Towards the Links of Cryptanalytic Methods on MPC/FHE/ZK-Friendly Symmetric-Key Primitives

Shiyao Chen[1,2,3], Chun Guo[3,4,6], Jian Guo[2], Li Liu[3,4], Meiqin Wang[3,4,5(✉)], Puwen Wei[3,4,5] and Zeyu Xu[3,4]

[1] Strategic Centre for Research in Privacy-Preserving Technologies and Systems, Nanyang Technological University, Singapore, Singapore
shiyao.chen@ntu.edu.sg

[2] Division of Mathematical Sciences, School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore, Singapore
guojian@ntu.edu.sg

[3] School of Cyber Science and Technology, Shandong University, Qingdao, China
{chun.guo,pwei,mqwang}@sdu.edu.cn,{sdu_liuli,xuzeyu}@mail.sdu.edu.cn

[4] Key Laboratory of Cryptologic Technology and Information Security, Ministry of Education, Shandong University, Qingdao, China

[5] Quan Cheng Laboratory, Jinan, China

[6] Shandong Research Institute of Industrial Technology, Jinan, China

Abstract. Symmetric-key primitives designed over the prime field $\mathbb{F}_p$ with odd characteristics, rather than the traditional $\mathbb{F}_2^n$, are becoming the most popular choice for MPC/FHE/ZK-protocols for better efficiencies. However, the security of $\mathbb{F}_p$ is less understood as there are highly nontrivial gaps when extending the cryptanalysis tools and experiences built on $\mathbb{F}_2^n$ in the past few decades to $\mathbb{F}_p$.

At CRYPTO 2015, Sun *et al.* established the links among impossible differential, zero-correlation linear, and integral cryptanalysis over $\mathbb{F}_2^n$ from the perspective of distinguishers. In this paper, following the definition of linear correlations over $\mathbb{F}_p$ by Baignères, Stern and Vaudenay at SAC 2007, we successfully establish comprehensive links over $\mathbb{F}_p$, by reproducing the proofs and offering alternatives when necessary. Interesting and important differences between $\mathbb{F}_p$ and $\mathbb{F}_2^n$ are observed.

- Zero-correlation linear hulls can not lead to integral distinguishers for some cases over $\mathbb{F}_p$, while this is always possible over $\mathbb{F}_2^n$ proven by Sun *et al.*.
- When the newly established links are applied to GMiMC, its impossible differential, zero-correlation linear hull and integral distinguishers can be increased by up to 3 rounds for most of the cases, and even to an *arbitrary number of rounds* for some special and limited cases, which only appeared in $\mathbb{F}_p$. It should be noted that all these distinguishers do not invalidate GMiMC's security claims.

The development of the theories over $\mathbb{F}_p$ behind these links, and properties identified (be it similar or different) will bring clearer and easier understanding of security of primitives in this emerging $\mathbb{F}_p$ field, which we believe will provide useful guides for future cryptanalysis and design.

Keywords: Symmetric-Key · Cryptanalysis · Proof · MPC/FHE/ZK-Friendly Primitives · Generalized Feistel · GMiMC

1 Introduction

With recent developments of practical cryptographic applications for advanced protocols, such as Multiparty Computation (MPC), Fully Homomorphic Encryption (FHE) and Zero-Knowledge proof (ZK), new criteria for symmetric-key primitives has been proposed. When choosing traditional standards like AES and SHA-3 as underlying primitives, it becomes the bottleneck of cryptographic computations. Naturally, a line of research of MPC/FHE/ZK-friendly symmetric-key primitives has been developed, including some MPC-friendly designs [AGR+16, AGP+19a, GLR+20, DGGK21], FHE-friendly designs [ARS+15, MJSC16, CCF+18, DEG+18, DGH+21, CIR22] and ZK-friendly designs [AD18, AAB+20, GKR+21, GKL+22, GHR+22, BBC+22].

Motivations. MPC/FHE/ZK have been one of the most popular lines of research in recent years, which brings researchers in different subareas of cryptography together. With many innovative and efficient symmetric-key primitives having been proposed, all these explorations may pose some potential threats to the security of these novel designs. Naturally, developing new collections of symmetric cryptanalytic tools over the prime field $\mathbb{F}_p$ is in urgent need, which could facilitate the design and cryptanalysis for researchers with a variety of backgrounds. The links of these symmetric cryptanalytic techniques have been important tools and well studied over $\mathbb{F}_2^n$ with many dedicated works [CV94, BN13, Lea11, BLNW12, SLR+15, BN14], among which linear cryptanalysis and its variants are the connections between these cryptanalytic methods to some extent. When considering the linear correlation over $\mathbb{F}_2^n$, parity-check is extensively used for its fast calculation, however, this is different for the correlation over $\mathbb{F}_p$, which is introduced and defined over a complex plane by Baignères *et al.* [BSV07] for better estimates but also more complicated. Therefore, full links among some popular symmetric cryptanalytic techniques over $\mathbb{F}_p$ are still missing. Beyne [Bey21] has recently provided new insights into linear cryptanalysis over abelian groups and generalized the link between zero-correlation and integral attacks, which are obtained by introducing a geometric approach. So, we wonder whether *comprehensive links among these symmetric cryptanalytic methods over* $\mathbb{F}_p$ can be built in a more popular way such as Bogdanov *et al.*'s work [BLNW12], and whether *different or similar properties between* $\mathbb{F}_p$ *and* $\mathbb{F}_2^n$ can be identified from the establishment of these links.

Contributions. In this paper, from the aspect of distinguishers, we establish the comprehensive links among impossible differential, zero-correlation linear and integral cryptanalysis over $\mathbb{F}_p$, for the very first time. From developments of the theories over $\mathbb{F}_p$ behind these links, similar and different properties are both identified, which will bring clearer and easier understanding of security of these MPC/FHE/ZK-friendly primitives. Then, as bonus and also applications, by using the proposed links, improved different types of distinguishers for GMiMC, a family of symmetric-key primitives proposed at ESORICS 2019 by Albrecht *et al.* [AGP+19b], are obtained. For the sake of simplicity, we will use DC, LC, IDC, ZC and INT to denote corresponding distinguishers or cryptanalytic methods for differential, linear, impossible differential, zero-correlation linear and integral cryptanalysis respectively in the rest of the paper. Our contributions are detailed as follows.

Comprehensive links among IDC, ZC and INT over $\mathbb{F}_p$*.* In Section 3, the links between IDC and ZC over $\mathbb{F}_p$ are established first. Then from the basic definition of linear correlation over $\mathbb{F}_p$, an alternative proof of the links between ZC and INT is presented, and we find that a ZC not always implies the existence of an INT over $\mathbb{F}_p$, however this is always possible over $\mathbb{F}_2^n$ proved in [SLR+15], which exhibits a difference between $\mathbb{F}_p$ and $\mathbb{F}_2^n$. Meanwhile, for another direction, we prove that an INT can lead to a ZC, if and only if it is a balanced integral distinguisher. Besides, to the best of our knowledge, there is no related works about the statistical complexity model of ZC over $\mathbb{F}_p$, and our proposed

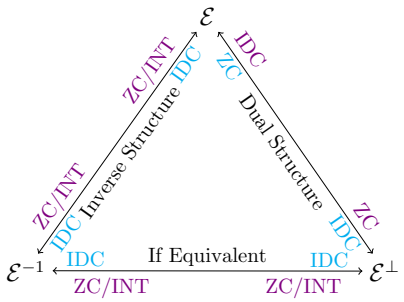


Figure 1: The IDC/ZC/INT transformations between the structure, its dual and inverse. (This figure abstracts the established comprehensive links, which could be intuitively explained by results on GMiMC$_{\texttt{Nyb}}$ in Section 4.3 later. To put it simply, with IDC of the structure $\mathcal{E}_{Nyb}$, it can be transformed into ZC/INT of the dual structure $\mathcal{E}_{Nyb}^{\perp}$, as $\mathcal{E}_{Nyb}^{\perp}$ is equivalent to the inverse structure $\mathcal{E}_{Nyb}^{-1}$, it finally will lead to ZC/INT back for $\mathcal{E}_{Nyb}$.)

links just provide a temporary solution when one wants to make use of ZC over $\mathbb{F}_p$ for further attack. With these formal treatments on the conditions and properties over $\mathbb{F}_p$, IDC and INT can be naturally linked, also including the difference between $\mathbb{F}_p$ and $\mathbb{F}_2^n$. When establishing these links, not only do we cover as many cipher constructions as possible, but also utilize the inverse structure $\mathcal{E}^{-1}$, together with the structure $\mathcal{E}$ and its dual $\mathcal{E}^{\perp}$ introduced in [BBW14, SLR+15], to explore more refined links with the potential equivalent relations of these structures. Finally, the comprehensive links among these symmetric cryptanalytic methods over $\mathbb{F}_p$ are established (abstracted in Figure 1), which facilitates to investigate more constructions but with less cryptanalysis efforts, and could be fundamental tools at hand for future design and cryptanalysis over $\mathbb{F}_p$.

Improvements of IDC, ZC and INT for all GMiMC constructions. To showcase the established links, we then apply to GMiMC and obtain improved distinguishers of IDC, ZC and INT for all GMiMC constructions in Section 4, summarized in Table 1 and Table 2. For unbalanced Feistel constructions GMiMC$_{\texttt{erf}}$ and GMiMC$_{\texttt{crf}}$, improved IDC, ZC and INT are obtained by using the equation-based method and the established links. With the condition $t \equiv 1 \bmod p$, *any number of rounds* of DC, LC, IDC, ZC and INT can be even constructed, from which a gap between $\mathbb{F}_p$ and $\mathbb{F}_2^n$ is also identified. Although this case is limited due to the large branches t and small field $\mathbb{F}_p$, it still fits some potential instantiations (e.g., two instances GMiMC$_{\texttt{erf}}$-$(p = 5, t = 86, r = 261)$[1] and GMiMC$_{\texttt{erf}}$-$(p = 17, t = 52, r = 160)$ provided in [AGP+19b, Table 6 and 7]) intended to compete with LowMC in post-quantum signatures, especially for the use-cases requiring full-data security where GMiMC will be used as a block cipher with 256-bit block/key size in Davies-Meyer construction to obtain a collision-resistant hash function. It should be noted that for low-data setting[2] in [CDG+17], except the LC with probability 1 that can be used to reduce 1-bit information of the preimage, these statistical distinguisher cannot be applied due to the limited data access where the chosen plaintext model is not suitable. As for two balanced Feistel GMiMC constructions, we reveal some underlying equivalent relations for both GMiMC$_{\texttt{Nyb}}$ and GMiMC$_{\texttt{mrf}}$, then by using our refined links, one-to-one correspondences between IDC, ZC and INT can be obtained. Finally combined with the equation-based methods, improved IDC, ZC and INT are also achieved for these two constructions.

[1] Here, compared to the most often used block size n for ciphers over $\mathbb{F}_2^n$, the block size of GMiMC over $\mathbb{F}_p$ is denoted by $t \cdot \log_2(p)$.

[2] GMiMC will be used as a one-way function f, where for the secret key x, its image $y = f(x)$ is published as the public key.

Table 1: Comparisons of different distinguishers of $\text{GMiMC}_{\texttt{erf}}$ and $\text{GMiMC}_{\texttt{crf}}$.

Ciphers	Type	Rounds	Remark	Time/Data Compl.*	Source
$\text{GMiMC}_{\texttt{erf}}$	IDC	$2t-2$	$\alpha_1, \beta_1 \neq 0$†	$O(p^{t-2})/O(p^{t-2})$	[AGP+19b]
		$3t-4$	$\alpha_1, \beta_1 \neq 0$ and $\alpha_1 \neq \beta_1$	$O(p^{t-2})/O(p^{t-2})$	[BCD+20]
		$3t-3$	$\alpha_1, \beta_1 \neq 0$	$O(p^{t-2})/O(p^{t-2})$	Sec 4.1.1
		$3t-1$	$\alpha_1 = \beta_1$ and $t \not\equiv 1 \bmod p$	$O(p^{t-1})/O(p^{t-1})$	Sec 4.1.2
		Arbitrary	$\alpha_1 = -\beta_1$ and $t \equiv 1 \bmod p$	$O(p^{t-1})/O(p^{t-1})$	Sec 4.1.3
	ZC	**$3t-3$**	$a_1, b_1 \neq 0$††	NA**	Sec 4.2.2
		$3t-1$	$a_1 = b_1$		Sec 4.2.3
		Arbitrary	$t \equiv 1 \bmod p$		Sec 4.2.1
	INT	$t + \lceil \log_d(t) \rceil$‡	Higher-order		[AGP+19b]
		$2t-3+\lfloor \log_d(p-2) \rfloor$	Block cipher usage	$O(p)/O(p)$	[BCD+20]
		$3t-3$		$O(p^{t-1})/O(p^{t-1})$	Sec 4.2.4
		$3t-4+\lfloor \log_d(p-2) \rfloor$	Hash function usage	$O(p)/O(p)$	[BCD+20]
		Arbitrary	$t \equiv 1 \bmod p$	$O(p)/O(p)$	Sec 4.2.4
	LC	$t-1$		$O(1)/O(1)$	[BCD+20]
		Arbitrary	$t \equiv 1 \bmod p$	$O(1)/O(1)$	Sec 4.2
	DC	$t \cdot (t+1) \cdot \lceil \frac{n}{2(n-1)} \rceil$‡‡	Truncated differential		[AGP+19b]
		$t^2 - t - 2$	Truncated differential	$O(p^{t-2})/O(p^{t-2})$	[BCD+20]
$\text{GMiMC}_{\texttt{crf}}$	IDC	**$3t-3$**	$\alpha_1, \beta_1 \neq 0$	$O(p^{t-2})/O(p^{t-2})$	Sec 4.2.4
		$3t-1$	$\alpha_1 = \beta_1$	$O(p^{t-1})/O(p^{t-1})$	Sec 4.2.4
		Arbitrary	$t \equiv 1 \bmod p$	$O(p^{t-1})/O(p^{t-1})$	Sec 4.2.4
	ZC	**$3t-3$**	$a_1, b_1 \neq 0$	NA	Sec 4.1.4
		$3t-1$	$a_1 = b_1$ and $t \not\equiv 1 \bmod p$		Sec 4.1.4
		Arbitrary	$a_1 = -b_1$ and $t \equiv 1 \bmod p$		Sec 4.1.4
	INT	$2t + \lceil \log_d(t) \rceil$‡	Higher-order		[AGP+19b]
		$3t-3$		$O(p^{t-1})/O(p^{t-1})$	Sec 4.1.4
	DC	$t \cdot (t+1) \cdot \lceil \frac{n}{2(n-1)} \rceil$‡‡	Truncated differential		[AGP+19b]
		$t^2 + t - 2$	Truncated differential	$O(p^{t-1})/O(p^{t-1})$	[BL22]
		Arbitrary	$t \equiv 1 \bmod p$	$O(1)/O(1)$	Sec 4.2

* Considering that previous IDC, INT of GMiMC are not provided with the success probability, for better comparisons, the corresponding time and data complexity are given here.
† α_1 and β_1 are related to the input and output differences respectively.
†† a_1 and b_1 are related to the input and output masks respectively.
** Not applicable since the corresponding statistical complexity theory of ZC over $\mathbb{F}_p$ is still missing.
‡ The bounds are roughly evaluated by algebraic degree and dimension of input (full codebook) by designers, where d is the degree of power map $S(x) \coloneqq x^d$ used in GMiMC constructions.
‡‡ The bounds are roughly evaluated for the iterative truncated differential (full codebook) by designers.

Table 2: Comparisons of different distinguishers of $\text{GMiMC}_{\texttt{Nyb}}$ and $\text{GMiMC}_{\texttt{mrf}}$.

Ciphers	Type	Rounds	Remark	Time/Data Complexity	Source
$\text{GMiMC}_{\texttt{Nyb}}^{\star}$	IDC	$2t-2$		$O(p^{t-2})/O(p^{t-2})$	[AGP+19b]
		$\mathbf{2t-1}$		$O(p^{t-2})/O(p^{t-2})$	Sec 4.3.1
		$\mathbf{2t+1}$	$\alpha_1=\beta_1$	$O(p^{t-1})/O(p^{t-1})$	Sec 4.3.2
	ZC	$2t-2$	Derived from IDC	NA	[AGP+19b]
		$\mathbf{2t-1}$			Sec 4.3.3
		$\mathbf{2t+1}$	$a_1=b_1$		Sec 4.3.3
	INT	$1+\lceil\log_d(t)\rceil$	Higher-order		[AGP+19b]
		$2t-2$	Derived from IDC	$O(p^{t-1})/O(p^{t-1})$	[AGP+19b]
		$\mathbf{2t-1}$		$O(p^{t-1})/O(p^{t-1})$	Sec 4.3.3
	DC	$3t$	Truncated differential		[AGP+19b]
$\text{GMiMC}_{\texttt{mrf}}^{\star}$	IDC	$2\Lambda(t)-2^{\star\star}$		$O(p^{t-2})/O(p^{t-2})$	[AGP+19b]
		$\mathbf{2\Lambda(t)-1}$	t is power-of-two	$O(p^{t-2})/O(p^{t-2})$	Sec 4.4.2
		$\mathbf{2\Lambda(t)+1}$	t is power-of-two and $\alpha_1=\beta_1$	$O(p^{t-1})/O(p^{t-1})$	Sec 4.4.1
	ZC	$2\Lambda(t)-2$	Derived from IDC	NA	[AGP+19b]
		$\mathbf{2\Lambda(t)-1}$	t is power-of-two		Sec 4.4.4
		$\mathbf{2\Lambda(t)+1}$	t is power-of-two and $a_1=b_1$		Sec 4.4.4
	INT	$\Lambda(t)+\lceil\log_d(t)\rceil$	Higher-order		[AGP+19b]
		$2\Lambda(t)-2$	Derived from IDC	$O(p^{t-1})/O(p^{t-1})$	[AGP+19b]
		$\mathbf{2\Lambda(t)-1}$	t is power-of-two	$O(p^{t-1})/O(p^{t-1})$	Sec 4.4.4
	DC	$3\Lambda(t)$	Truncated differential		[AGP+19b]

* For $\text{GMiMC}_{\texttt{Nyb}}$ and $\text{GMiMC}_{\texttt{mrf}}$, the number of branch t is even and $t \geq 4$.
** $\Lambda(t) = 2\lceil\log_2(t)\rceil$, the minimum number of rounds to reach full diffusion.

Comparisons to previous works on GMiMC. There are some prior dedicated cryptanalysis results [Bon19, BCD+20, BL22] on GMiMC. In [Bon19], Bonnetain observed that there exists special slide attacks on GMiMC with key size $\log_2 p$, that is the univariate case which will be introduced later in Section 2.2. Thus, due to the weaknesses found for GMiMC univariate case, *we only consider GMiMC permutations or block ciphers with full key size (multivariate case) in this paper.* Later, Beyne *et al.* [BCD+20] focused on GMiMC permutations adopted in sponge-based construction where no key materials are involved, and they finally proposed improved INT, IDC and DC for $\text{GMiMC}_{\texttt{erf}}$. Recently, Beyne *et al.* [BL22] also provided elaborate truncated differential cryptanalysis on $\text{GMiMC}_{\texttt{crf}}$. Some of these results are listed in Table 1, and below we detail some comparisons.

- **INT.** As said in [AGP+19b], "...attacks that do not depend on the round function, become competitive. Still, for practical use cases we show that a high number of branches can be meaningful...", compared to the dedicated degree-based method, our link-based method covers both keyed and unkeyed settings and is independent of the round function (i.e. the power map x^d and the field $\mathbb{F}_p$), which could reveal underlying structural properties. While in [BCD+20], an INT with $3t-4+log_d(p-2)$ rounds (d is the degree of the power map) is constructed with the dedicated degree-based method and only works for the permutation used in hash function, which will

reduce to $2t-3+log_d(p-2)$ rounds for block cipher usage due to subkeys added for the first $2t-2$ rounds. Thus, our method provides a convenient way to derive the structural distinguishers that can capture the underlying structural properties for the target construction, which will be basic and convenient tools for both designers and cryptanalysts.

- **DC.** We note that truncated differential cryptanalysis seems more powerful than other statistical attacks on GMiMC to date. However, on the one hand, we would like to emphasize that the comprehensive links established in this paper are among IDC, ZC and INT over $\mathbb{F}_p$, to apply our links, it is mainly expected to improve these three kinds of distinguishers first, which could show the convenience and effectiveness of our method. On the other hand, from the view of designers, all attack vectors should be taken into consideration, thus it is still necessary to explore bounds of different kinds of distinguishers, such as IDC, INT improvements provided in [BCD+20]. Besides, to establish these refined links over $\mathbb{F}_p$, we focus more on different kinds of constructions and structural distinguishers (i.e. IDC, ZC and INT). Hence, mounting preimage/collision[3] or key-recovery attacks on some concrete ciphers are not our goals, which could be left for future works.

2 Preliminaries

In this section, we give some preliminaries, including the differential, linear and integral properties, and structure, dual structure and inverse structure of symmetric-key primitives, which will be used in the proofs of the links over $\mathbb{F}_p$ in Section 3, and we mainly consider the prime field $\mathbb{F}_p$ with odd characteristic in this paper. For the later applications in Section 4, GMiMC ciphers are also briefly introduced.

2.1 Differential, Linear and Integral Cryptanalysis over $\mathbb{F}_p$

Differential cryptanalysis over $\mathbb{F}_p$: The differential probability of the function F over $\mathbb{F}_p^t$ can be easily generalized as below

$$prob_F(\alpha,\beta)=\frac{|x:F(x)-F(x-\alpha)=\beta|}{p^t},$$

where $\alpha,\beta,x\in\mathbb{F}_p^t$. Considering the commonly used key addition in symmetric-key primitives over $\mathbb{F}_p$, the modular subtraction difference is adopted here.

Linear cryptanalysis over $\mathbb{F}_p$: Baignères *et al.* [BSV07] have developed the correlation analysis of primitives that operate on the prime field, which has been recently used to evaluate the security of Ciminion [DGGK21] against linear attacks. The core idea is that a character is an additive homomorphism from $\mathbb{F}_p^t$ into $S_p=\{z\in\mathbb{C}:z^p=1\}$ and any character is of the form

$$\chi_u(x)=e^{\frac{2\pi i}{p}u^T\cdot x},$$

where $x,u\in\mathbb{F}_p^t$. The following definition of the correlation over $\mathbb{F}_p$ is introduced.

Definition 1 (Correlation over $\mathbb{F}_p$ [BSV07])**.** Given a function $F:\mathbb{F}_p^t\to\mathbb{F}_p^s$, for a linear mask pair (u,v), where $u\in\mathbb{F}_p^t$ and $v\in\mathbb{F}_p^s$, then the correlation of the linear approximation (u,v) of F is defined as

$$cor_F(u,v)=cor(u^T\cdot x-v^T\cdot F(x))=\frac{1}{p^t}\sum_{x\in\mathbb{F}_p^t}\chi_u(x)\overline{\chi_v(F(x))}=\frac{1}{p^t}\sum_{x\in\mathbb{F}_p^t}e^{\frac{2\pi i}{p}(u^T\cdot x-v^T\cdot F(x))}.$$

[3] IDC, ZC and INT are not suitable for preimage/collision attacks.

According to this definition, the correlation over $\mathbb{F}_p$ can be evaluated by a complex number with its norm located in $[0, 1]$, and the general linear probability can be defined as follows.

Definition 2 (Linear probability over $\mathbb{F}_p$ [BSV07])**.** $lprob_F(u, v) = |cor_F(u, v)|^2$.

Zero-correlation linear hull has been introduced by Bogdanov and Rijmen [BR14], based on the linear correlation defined over $\mathbb{F}_p$, it can be naturally generalized to $\mathbb{F}_p$.

Definition 3 (Zero-correlation linear hull over $\mathbb{F}_p$)**.** Given a function $F : \mathbb{F}_p^t \to \mathbb{F}_p^s$, for the mask pair (u, v) where $u \in \mathbb{F}_p^t$ and $v \in \mathbb{F}_p^s$, then (u, v) is called a zero-correlation linear hull of F, if and only if $cor_F(u, v) = 0$.

Given the definitions above, the propagations of linear mask over $\mathbb{F}_p$ of some basic operations can be obtained. Similarly to $\mathbb{F}_2^n$, for the branching operation $x \to (x, x)$ where $x \in \mathbb{F}_p$, for the linear masks $a \to (b, c)$, it must have $a = b + c$; For the addition operation, $x + y = z$ where $x, y, z \in \mathbb{F}_p$, for the linear masks $(a, b) \to c$, it must have $a = b = c$. For more detailed proofs and other operations, we refer the reader to [DGGK21, Appendix C.2]. Furthermore, the following properties over $\mathbb{F}_p$ can be deduced.

Proposition 1. *For any fixed non-zero $a \in \mathbb{F}_p$, $cor(a \cdot x) = 0$.*

Proof. As a is non-zero, for the complex number $e^{\frac{2\pi i}{p}a} = \cos(\frac{2\pi}{p}a) + \sin(\frac{2\pi}{p}a)i$, we have $e^{\frac{2\pi i}{p}a} \neq 1$. Considering the complex multiplication,

$$e^{\theta_0 i} \times e^{\theta_1 i} = (\cos\theta_0 + i\sin\theta_0) \times (\cos\theta_1 + i\sin\theta_1) = \cos(\theta_0 + \theta_1) + \sin(\theta_0 + \theta_1) = e^{(\theta_0+\theta_1)i}.$$

Then according to the property of the geometric sequence, it has the following

$$cor(a \cdot x) = \frac{1}{p}\sum_{x\in\mathbb{F}_p} e^{\frac{2\pi i}{p}ax} = \frac{1}{p}(e^{\frac{2\pi i}{p}0} + e^{\frac{2\pi i}{p}a} + \cdots e^{\frac{2\pi i}{p}(p-1)a}) = \frac{1 - e^{2\pi ai}}{p(1 - e^{\frac{2\pi i}{p}a})}.$$

As $e^{2\pi ai} = 1$, then $cor(a \cdot x) = 0$. □

Proposition 1 can be directly generalized to dimension t as follows.

Corollary 1. *For any fixed non-zero $a \in \mathbb{F}_p^t$, $cor(a^T \cdot x) = 0$.*

Proof. Let $a = (a_1, \cdots, a_t) \in \mathbb{F}_p^t$ and $x = (x_1, \cdots, x_t) \in \mathbb{F}_p^t$, then we have

$$\begin{aligned} cor(a^T \cdot x) &= \frac{1}{p^t}\sum_{x\in\mathbb{F}_p^t} e^{\frac{2\pi i}{p}a^T\cdot x} = \frac{1}{p^t}\sum_{x\in\mathbb{F}_p^t} e^{\frac{2\pi i}{p}(a_1x_1+\cdots+a_tx_t)} \\ &= \prod_{1\le i\le t}(\frac{1}{p}\sum_{x_i\in\mathbb{F}_p} e^{\frac{2\pi i}{p}a_ix_i}) = \prod_{1\le i\le t} cor(a_i^T \cdot x_i). \end{aligned}$$

Due to non-zero a, there must be at least one $a_i \neq 0$. According to Proposition 1, it will lead to $cor(a_i^T \cdot x_i) = 0$, which ends our proof. □

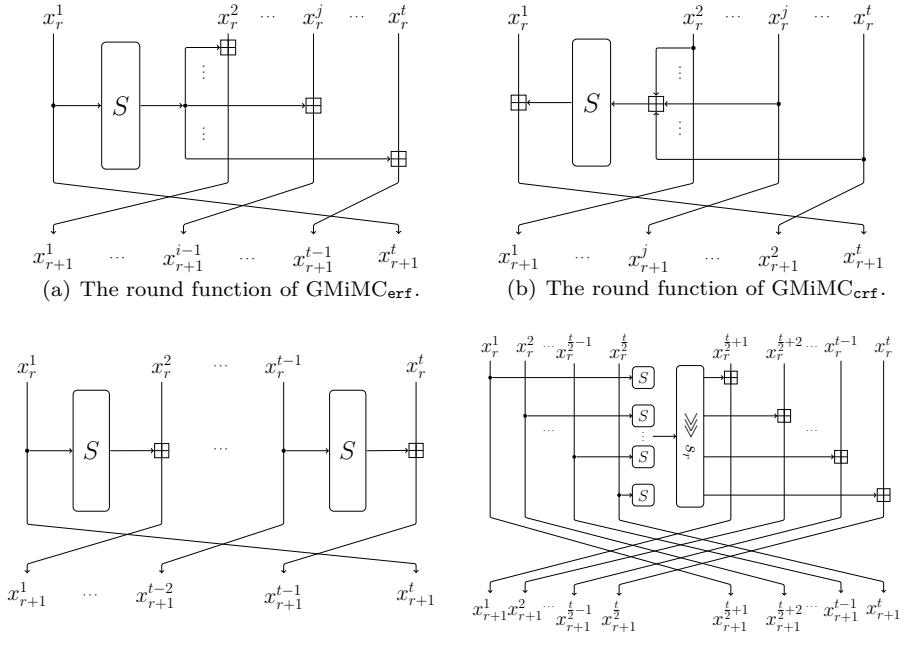
Integral cryptanalysis over $\mathbb{F}_p$: The notion of integral attacks has been introduced by Knudsen and Wagner [KW02], which captures several variants including high-order differential attack [Lai94] and saturation attack [Luc01]. Higher-order differentials over $\mathbb{F}_p$ can also make use of a generalized notion of differentiation as analyzed by Lai in [Lai94] (also refer to [AP11]). Recently, Beyne *et al.* show that the same technique can be used over $\mathbb{F}_p$, which further can be extended to multiplicative subgroups (see [Bey21, Proposition 1, Corollary 1, Proposition 2]), and this kind of degree-based integral distinguisher may not have the balanced property defined as below.

Definition 4 (Balanced property over $\mathbb{F}_p$)**.** Given a function $F : \mathbb{F}_p^t \to \mathbb{F}_p^s$, let A be a subspace of $\mathbb{F}_p^t$, if the size of the set $F_A(y) \triangleq \{x \in A | F(x) = y\}$ is independent of $y \in \mathbb{F}_p^s$, we say F is balanced on A.

It can be observed that if F is balanced on A, then it has the *balanced integral (zero-sum) property*, i.e. $\sum_{x \in A} F(x) = 0$. It should be noted that in this paper we will focus more on this kind of balanced integral distinguisher, which could reveal more underlying structure properties of the ciphers.

2.2 Specifications of GMiMC

GMiMC is a family of symmetric-key primitives designed by Albrecht *et al.* [AGP+19b] based on several generalized (unbalanced and balanced) Feistel networks using power maps $S(x) \coloneqq x^d$ as the non-linear component of the round function, e.g., GMiMC$_{\texttt{erf}}$ with expanding round function, GMiMC$_{\texttt{crf}}$ with contracting round function, GMiMC$_{\texttt{Nyb}}$ with Nyberg's GFN structure and GMiMC$_{\texttt{mrf}}$ with a new structure named Multi-Rotating by the designers, where different rotation parameters s_r are chosen for different rounds to change the positions of these S-boxes (please see Figure 2(d)). As these permutations of GMiMC can be used to construct both hash functions and block ciphers, we just depict the round functions of the corresponding permutations in Figure 2. For block cipher usage, GMiMC block cipher supports two key sizes: univariate case $\log_2(p)$ and multivariate case $t \cdot \log_2(p)$. The rounds are numbered starting from 1, and the branches are numbered from 1 to t where Branch 1 is the leftmost branch. For example, the state of Branch 1 and round r of GMiMC$_{\texttt{erf}}$ is represented by x_r^1 in Figure 2(a) and $x_r^1 \in \mathbb{F}_p$ for the chosen prime p. Thus, we denote the concrete instance of GMiMC permutation by GMiMC-(p, t, R) where R is the total number of rounds. For more details of GMiMC, we refer the reader to the design paper [AGP+19b].



(a) The round function of GMiMC$_{\texttt{erf}}$.

(b) The round function of GMiMC$_{\texttt{crf}}$.

(c) The round function of GMiMC$_{\texttt{Nyb}}$.

(d) The round function of GMiMC$_{\texttt{mrf}}$.

Figure 2: Four generalized Feistel networks adopted in GMiMC.

2.3 Structure, Dual Structure and Inverse Structure of Symmetric-Key Primitives over $\mathbb{F}_p$

The structure and dual structure of block ciphers over $\mathbb{F}_2^n$ have been introduced in [BBW14] to obtain the *equivalence* between different structures, which are also used in [SLR+15]. Together with the inverse structure utilized in this paper, we adapt these definitions to symmetric-key primitives over $\mathbb{F}_p$.

Definition 5. Let $E : \mathbb{F}_p^t \to \mathbb{F}_p^t$ be a permutation, which can be decomposed into S-boxes (the non-linear part) and linear transformations (the linear part). The internal state of E is represented by t elements of $\mathbb{F}_p$.

(1) A structure $\mathcal{E}^E$ over $\mathbb{F}_p^t$ is defined as a set of primitives, which is exactly same as E except that S-boxes can take all possible transformations on corresponding domains.

(2) Let $a, b \in \mathbb{F}_p^t$. If for all $E' \in \mathcal{E}^E$, $a \to b$ is an impossible differential (zero correlation linear hull) of E', then $a \to b$ is called an impossible differential (zero correlation linear hull) of $\mathcal{E}^E$.

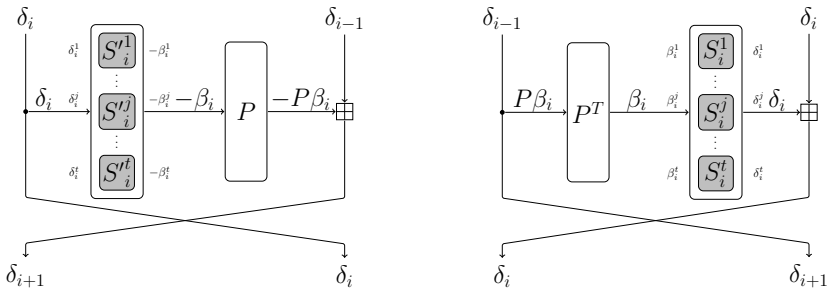
(3) Let $a, b \in \mathbb{F}_p^t$. If for all $E' \in \mathcal{E}^E$, $a \to b$ is a differential (linear) trail of E' with differential (linear) probability 1, then $a \to b$ is called a Prob-one differential (linear) trail of $\mathcal{E}^E$.

If E using bijective S-boxes, then S-boxes adopted in $\mathcal{E}^E$ should also be bijective. However, if S-boxes used in E are not limited to bijective, then $\mathcal{E}^E$ is defined as a set of the permutation E' which is exactly same as E except that S-boxes can take all possible transformations. Now, we adapt the definition of dual structure in [SLR+15] to $\mathbb{F}_p$ and cover the generalized Feistel structure in [BMT13].

Definition 6. We give the dual structure of classical balanced Feistel structure, generalized Feistel structure, SPN structure and two unbalanced Feistel structures as below.

- Let $\mathcal{F}_{SP}$ be a Feistel structure with SP-type round function, the state of which first passes the non-linear layer S then the linear transformation P. By abuse of notation, we also use P as the matrix representation for the linear layer in the rest of the paper, whose transpose and inverse are P^T and P^{-1} respectively. Let σ be the operation that exchanges the left and right halves of a state. Then the dual structure $\mathcal{F}_{SP}^{\perp}$ of $\mathcal{F}_{SP}$ is defined as $\sigma \circ \mathcal{F}_{P^T S} \circ \sigma$, the state of which passes σ operation, the linear transformation P^T, the non-linear S and σ operation.
- Let $\mathcal{GF}_{FP}$ be a Generalized Feistel structure (including the Extended Generalized Feistel structure) defined in [BMT13], where F is the non-linear part of the round function and adopts the matrix representation used in [BMT13], P is the linear transformation. Then the dual structure $\mathcal{GF}_{FP}^{\perp}$ of $\mathcal{GF}_{FP}$ is defined as $\mathcal{GF}_{F^T(P^{-1})^T}$.
- Let $\mathcal{E}_{SP}$ be an SPN structure with the non-linear S first and followed by the linear transformation P. Then the dual structure $\mathcal{E}_{SP}^{\perp}$ is defined as $\mathcal{E}_{S(P^{-1})^T}$.
- Let $\mathcal{E}_{erf}$ be a structure $\mathcal{E}^{\mathrm{GMiMC_{erf}}}$ and $\mathcal{E}_{crf}$ be a structure $\mathcal{E}^{\mathrm{GMiMC_{crf}}}$. Then structures $\mathcal{E}_{erf}$ and $\mathcal{E}_{crf}$ are dual with each other.

Since we do not consider the details of the S-box, by abuse of notation, S in structures is just to signify the order of the S-box layer and it is not a concrete S-Box layer. A demonstration of the structure (see Figure 3(a)) and its dual structure (see Figure 3(b)) are given for the classical Feistel structure. It should be noted that $\mathrm{GMiMC_{Nyb}}$ and $\mathrm{GMiMC_{mrf}}$ are covered by $\mathcal{GF}_{FP}$ and $\mathcal{F}_{SP}$ respectively, for the sake of simplicity, notations $\mathcal{E}_{Nyb}$, $\mathcal{E}_{mrf}$ and their dual structures will also be used in the rest of the paper. The inverse structure is introduced as follows.



(a) Differential trail of $\mathcal{F}_{SP}$. (b) Linear trail of $\mathcal{F}^{\perp}_{SP}$.

Figure 3: Different and linear trails of $\mathcal{F}_{SP}$ and $\mathcal{F}^{\perp}_{SP}$.

Definition 7. We give the inverse structure of classical balanced Feistel structure, generalized Feistel structure, SPN structure and two unbalanced Feistel structures as below.

- Let $\mathcal{F}_{SP}$ be a Feistel structure with SP-type round function, and let the primitive representation of the linear transformation be P. Let σ be the operation that exchanges the left and right halves of a state. Then the inverse structure $\mathcal{F}^{-1}_{SP}$ of $\mathcal{F}_{SP}$ is defined as $\sigma \circ \mathcal{F}_{SP} \circ \sigma$.

- Let $\mathcal{GF}_{FP}$ be a Generalized Feistel structure (including the Extended Generalized Feistel structure) defined in [BMT13], where F is the non-linear part of the round function and followed by the linear transformation P. Then the inverse structure $\mathcal{GF}^{-1}_{FP}$ of $\mathcal{GF}_{FP}$ is defined as $\mathcal{GF}_{P^{-1}F^{-1}}$.

- Let $\mathcal{E}_{SP}$ be an SPN structure with the non-linear S first and followed by the linear transformation P. Then the inverse structure $\mathcal{E}^{-1}_{SP}$ is defined as $\mathcal{E}_{P^{-1}S}$.

- Let $\mathcal{E}_{erf}$ be a structure $\mathcal{E}^{\mathrm{GMiMC_{erf}}}$ and $\mathcal{E}_{crf}$ be a structure $\mathcal{E}^{\mathrm{GMiMC_{crf}}}$. Then the corresponding inverse structures are $\mathcal{E}^{-1}_{erf}$ and $\mathcal{E}^{-1}_{crf}$ respectively.

3 Links among Impossible differential, Zero-correlation linear and Integral Cryptanalysis over $\mathbb{F}_p$

We start by giving the links between IDC and ZC over $\mathbb{F}_p$, and more refined links are obtained by covering more constructions and some equivalent relations. Then, we build the links between ZC and INT over $\mathbb{F}_p$, from which differences between $\mathbb{F}_p$ and $\mathbb{F}_2^n$ are observed. Finally, with the bridge between previous links (from IDC to ZC and ZC to INT), we provide the links between IDC and INT over $\mathbb{F}_p$. It should be noted that some basic properties of differential and linear over $\mathbb{F}_p$ are employed in a nontrivial way, and unlike the analogue on $\mathbb{F}_2^n$, addition on $\mathbb{F}_p$ is not involutional and the only nontrivial linear subspace over $\mathbb{F}_p$ is itself. Due to these, ZC/IDC does not always imply INT, and we need to characterize the sufficient conditions, which exhibits the essential difference of the links between $\mathbb{F}_p$ and $\mathbb{F}_2^n$.

As we consider the structure, dual structure and inverse structure, if not specified, the S-box adopted in these structures will be regarded as the *ideal S-box*, that is, any active input difference (mask) will lead to any active output difference (mask) and inactive input difference (mask) only produces inactive output difference (mask).

3.1 Links between IDC and ZC over $\mathbb{F}_p$

Similar to the proofs by Sun *et al.* [SLR+15], the transformations over $\mathbb{F}_p$ between IDC and ZC are proved in following two lemmas from two directions, which are also extended to more constructions and structures.

Lemma 1. *For a linear hull $(\delta_0, \delta_1) \to (\delta_r, \delta_{r+1})$, if there exists $E \in \mathcal{F}_{SP}^{\perp}$ such that*

$$cor((\delta_0, \delta_1) \cdot x - (\delta_r, \delta_{r+1}) \cdot E(x)) \neq 0,$$

then there exists $E' \in \mathcal{F}_{SP}$ such that

$$prob_{E'}((\delta_1, \delta_0), (\delta_{r+1}, \delta_r)) > 0.$$

Proof. As $(\delta_0, \delta_1) \to (\delta_r, \delta_{r+1})$ is a linear hull of some $E \in \mathcal{F}_{SP}^{\perp}$ with non-zero correlation, also see Figure 3(b). Then, according to definitions of linear probability of linear characteristic and linear hull over $\mathbb{F}_p$ [BSV07, Section 3.2], there must be a linear characteristic with non-zero correlation

$$(\delta_0, \delta_1) \to \cdots \to (\delta_i, \delta_{i+1}) \cdots \to (\delta_r, \delta_{r+1}),$$

where the input of the round function can be divided into t pieces of $\mathbb{F}_p$ elements, that is $\delta_i \in \mathbb{F}_p^t$. Considering this linear characteristic, the output mask of the non-linear layer $S_i = (S_i^1, \cdots, S_i^t)$ is $\delta_i = (\delta_i^1, \cdots, \delta_i^t) \in \mathbb{F}_p^t$. The input mask of S_i is denoted by $\beta_i = (\beta_i^1, \cdots, \beta_i^t) \in \mathbb{F}_p^t$. While for the linear layer P^T, denoting its input mask is $\gamma_i \in \mathbb{F}_p^t$ and input value is $x_i \in \mathbb{F}_p^t$, then we have

$$cor(\gamma_i^T \cdot x_i - \beta_i^T \cdot (P^T \cdot x_i)) = cor((\gamma_i^T - \beta_i^T P^T) \cdot x_i) = cor((\gamma_i - P\beta_i)^T \cdot x_i).$$

If $\gamma_i \neq P\beta_i$, according to Corollary 1, $cor((\gamma_i - P\beta_i)^T \cdot x_i) = 0$, which is contradicted with the non-zero correlation of this linear characteristic. Thus, $\delta_{i-1} = \delta_{i+1} + \gamma_i = \delta_{i+1} + P\beta_i$ must hold. Now focusing on the dual structure, for any plaintext (x_L, x_R), we can construct an r-round cipher $E_r \in \mathcal{F}_{SP}$, such that $E_r(x_L, x_R) - E_r(x_L - \delta_1, x_R - \delta_0) = (\delta_{r+1}, \delta_r)$.

When $r = 1$, for $j \in \{1, ..., t\}$: if $\delta_1^j = 0$, we can define S'^j_1 as any possible transformation over $\mathbb{F}_p$, and if $\delta_1^j \neq 0$, we can define the following

$$S'^j_1(x_L^j) = x_L^j \text{ and } S'^j_1(x_L^j - \delta_1^j) = x_L^j + \beta_1^j.$$

Then for $E_1 \in \mathcal{F}_{SP}$ which adopts such S-boxes, there will be

$$E_1(x_L, x_R) - E_1(x_L - \delta_1, x_R - \delta_0) = (\delta_0 + (-P\beta_1), \delta_1) = (\delta_2, \delta_1).^4$$

Suppose that we have constructed E_{r-1} such that

$$E_{r-1}(x_L, x_R) - E_{r-1}(x_L - \delta_1, x_R - \delta_0) = (\delta_r, \delta_{r-1}),$$

and let $(y_L, y_R) = (y_L^1, \cdots, y_L^t, y_R^1, \cdots, y_R^t)$ denote the output of $E_{r-1}(x_L, x_R)$. Then in the r-th round, if $\delta_r^j = 0$, we can define S'^j_r as any possible transformation over $\mathbb{F}_p$, otherwise, define S'^j_r as follows

$$S'^j_r(y_L^j) = y_L^j \text{ and } S'^j_r(y_L^j - \delta_r^j) = y_L^j + \beta_r^j.$$

Therefore, $E_r(x_L, x_R) - E_r(x_L - \delta_1, x_R - \delta_0) = (\delta_{r-1} - P\beta_r, \delta_r) = (\delta_{r+1}, \delta_r)$. □

[4] We should be careful about the sign (i.e. addition and subtraction), which is different over $\mathbb{F}_p$.

Lemma 2. *For a differential $(\delta_1, \delta_0) \to (\delta_{r+1}, \delta_r)$, if there exists $E \in \mathcal{F}_{SP}$ such that*

$$prob_E((\delta_1, \delta_0), (\delta_{r+1}, \delta_r)) > 0,$$

then there exists $E' \in \mathcal{F}_{SP}^{\perp}$ such that

$$cor((\delta_0, \delta_1) \cdot x - (\delta_r, \delta_{r+1}) \cdot E'(x)) \neq 0.$$

Proof. As $(\delta_1, \delta_0) \to (\delta_{r+1}, \delta_r)$ is a differential of some $E \in \mathcal{F}_{SP}$ with non-zero differential probability, also see Figure 3(a), then there must exist a differential characteristic with non-zero probability, denoted as

$$(\delta_1, \delta_0) \to \cdots \to (\delta_{i+1}, \delta_i) \cdots \to (\delta_{r+1}, \delta_r),$$

where $\delta_i \in \mathbb{F}_p^t$. For this differential characteristic, the input difference of the non-linear layer $S_i' = ({S'}_i^1, \cdots, {S'}_i^t)$ is $\delta_i = (\delta_i^1, \cdots, \delta_i^t) \in \mathbb{F}_p^t$. The output difference of S_i' is denoted by $-\beta_i = (-\beta_i^1, \cdots, -\beta_i^t) \in \mathbb{F}_p^t$, then $\delta_{i-1} - P\beta_i = \delta_{i+1}$.

Considering the following fact: for mask pair (β_i^j, δ_i^j), where $\delta_i^j \neq 0$, there always exists an element $a_i^j \in \mathbb{F}_p$ such that $\beta_i^j = a_i^j \delta_i^j$, then for $S_i^j(x) = a_i^j x$, we have $cor((\beta_i^j)^T \cdot x - (\delta_i^j)^T \cdot S_i^j(x)) = cor((\beta_i^j - a_i^j \delta_i^j)^T \cdot x) = 1$.

Now for the dual structure, we construct an r-round cipher $E_r \in \mathcal{F}_{SP}^{\perp}$ such that $cor((\delta_0, \delta_1) \cdot x - (\delta_r, \delta_{r+1}) \cdot E_r(x)) \neq 0$. If $r = 1$, let $S_1^j(x) = a_1^j x$ for $\delta_1^j \neq 0$ and any linear transformation over $\mathbb{F}_p$ otherwise. Then all operations in $E_1 \in \mathcal{F}_{SP}^{\perp}$ are linear over $\mathbb{F}_p$, which implies that there exists an affine transformation $L_1(x) = A_1 x + B_1$, where $x \in \mathbb{F}_p^{2t}$, A_1 is a $2t \times 2t$ matrix over $\mathbb{F}_p$ and B_1 is a $2t$-dimensional vector over $\mathbb{F}_p$, such that $E_1(x) = L_1 x$ and with

$$cor((\delta_0, \delta_1) \cdot x - (\delta_1, \delta_2) \cdot E_1(x)) = 1.$$

Assume that we have $E_{r-1}(x) = L_{r-1}x = A_{r-1}x + B_{r-1}$ where A_{r-1} is a $2t \times 2t$ matrix over $\mathbb{F}_p$ and B_{r-1} is a $2t$-dimensional vector over $\mathbb{F}_p$ such that

$$cor((\delta_0, \delta_1) \cdot x - (\delta_{r-1}, \delta_r) \cdot E_{r-1}(x)) = 1.$$

We then define $S_{r,j}(x)$ in the r-th and have $E_r(x) = L_r x = A_r x + B_r$ where A_r is a $2t \times 2t$ matrix over $\mathbb{F}_p$ and B_r is a $2t$-dimensional vector over $\mathbb{F}_p$ such that

$$cor((\delta_0, \delta_1) \cdot x - (\delta_r, \delta_{r+1}) \cdot E_r(x)) = 1.$$

Thus, we have $cor((\delta_0, \delta_1) \cdot x - (\delta_r, \delta_{r+1}) \cdot E_r(x)) \neq 0$. □

Theorem 1. *Over $\mathbb{F}_p$, $a \to b$ is an impossible differential of $\mathcal{F}_{SP}$ if and only if it is a zero-correlation linear hull of $\mathcal{F}_{SP}^{\perp}$.*

Proof. We consider the following two parts.

(1) Assume $a \to b$ is an impossible differential of $\mathcal{F}_{SP}$, if it is not a zero-correlation linear hull of $\mathcal{F}_{SP}^{\perp}$. Then, according to Lemma 1, there must be some $E' \in \mathcal{F}_{SP}$ such that $prob_{E'}((\delta_1, \delta_0) \to (\delta_{r+1}, \delta_r)) > 0$, which contradicts that $a \to b$ is an impossible differential of $\mathcal{F}_{SP}$.

(2) Similarly, assume $a \to b$ is a zero-correlation linear hull of $\mathcal{F}_{SP}^{\perp}$, if it is not an impossible differential of $\mathcal{F}_{SP}$. Then, according to Lemma 2, then must be $E' \in \mathcal{F}_{SP}^{\perp}$ such that $cor((\delta_0, \delta_1) \cdot x - (\delta_r, \delta_{r+1}) \cdot E'(x)) \neq 0$, which contradicts that $a \to b$ is a zero-correlation linear hull of $\mathcal{F}_{SP}^{\perp}$.

As claimed. □

Note that to focus more on the proofs of the structures in Lemma 1 and Lemma 2, we do not limit the constructed S-box to be a bijective one. If the adopted S-box is bijective, these two lemmas still hold. In Lemma 1, for a bijective S-box, if the correlation is non-zero, then the output mask $\delta_i^j \neq 0$ implies the input mask $\beta_i^j \neq 0$. We have the following S-box ${S'}_r^j$ to satisfy the bijective condition and difference transitions.

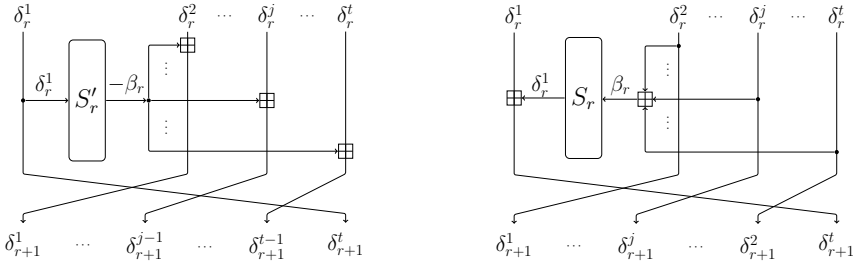
$$ {S'}_r^j(x) = \begin{cases} x_r^j - \delta_r^j, & x = x_r^j + \beta_r^j, \\ x_r^j + \beta_r^j, & x = x_r^j - \delta_r^j, \\ x, & \text{others.} \end{cases} $$

While in Lemma 2, for a bijective S-box, if the differential probability is non-zero, then the input difference $\delta_i^j \neq 0$ implies the output difference $-\beta_i^j \neq 0$. Thus, we can define the S-box $S_r^j(x) = a_r^j x$ ($a_r^j \neq 0$ and $-\beta_i^j = a_r^j \delta_r^j$), which satisfies the bijective condition and linear mask propagations.

For the above proofs of the classical Feistel network with the SP-type round function, an abstract of the S-box layer S and the matrix representation of linear layer P are used. When considering the proofs of this kind of SP-type round function for other constructions, similar theorems can be obtained as follows for the SPN construction and Generalized Feistel Networks introduced in [BMT13], where generic matrix representations for both non-linear and linear layers have been proposed.

Theorem 2. *Over $\mathbb{F}_p$, $a \to b$ is an impossible differential of $\mathcal{E}_{SP}$ if and only if it is a zero-correlation linear hull of $\mathcal{E}_{SP}^{\perp}$.*

Theorem 3. *Over $\mathbb{F}_p$, $a \to b$ is an impossible differential of $\mathcal{GF}_{FP}$ if and only if it is a zero-correlation linear hull of $\mathcal{GF}_{FP}^{\perp}$.*



(a) Differential trail of $\mathcal{E}_{erf}$.

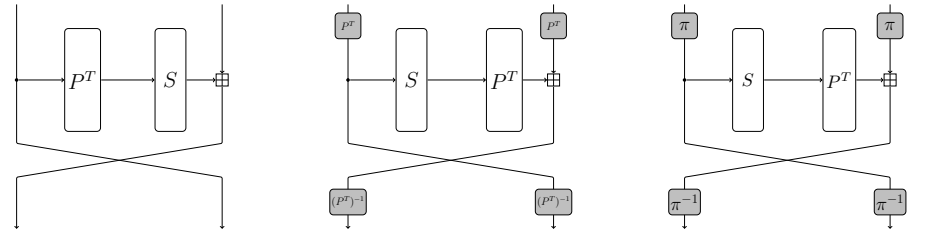
(b) Linear trail of $\mathcal{E}_{crf}$.

Figure 4: Differential and linear trails of $\mathcal{E}_{erf}$ and $\mathcal{E}_{crf}$.

Still similar to the proofs of Lemma 1 and Lemma 2, we can prove the following theorem for the structure $\mathcal{E}_{erf}$ and its dual $\mathcal{E}_{crf}$ for GMiMC, see Figure 4(a) and Figure 4(b). These two unbalanced Feistel structures are not covered by the definitions in [BMT13], and the detailed proof of Theorem 4 is provided in Appendix A.

Theorem 4. *$a \to b$ is an impossible differential (zero-correlation linear hull) of $\mathcal{E}_{erf}$ if and only if it is a zero-correlation linear hull (impossible differential) of $\mathcal{E}_{crf}$.*

Corollary 2. *Let $\mathcal{F}_{SP}$ be a Feistel structure with SP-type round function, and the linear transformation be P. If P is invertible, an impossible differential of $\mathcal{F}_{SP}$ is equivalent to a zero-correlation linear hull of $\mathcal{F}_{SP^T}$.*



(a) Structure of $\mathcal{F}_{P^TS}$. (b) Equivalent structure of $\mathcal{F}_{SP^T}$. (c) $\mathcal{F}_{SP}$ with $P = \pi^{-1}P^T\pi$.

Figure 5: Structures of $\mathcal{F}_{P^TS}$, $\mathcal{F}_{SP^T}$ and $\mathcal{F}_{SP}$.

Proof. As P is invertible, according to the definition of equivalent structures given in [LLF05], which are depicted in Figure 5(a) and Figure 5(b), we have

$$\mathcal{F}_{P^TS} = ((P^T)^{-1}, (P^T)^{-1}) \circ \mathcal{F}_{SP^T} \circ (P^T, P^T).$$

Thus, combining with Theorem 1, we can end the proof. □

Corollary 3. *For a Feistel structure $\mathcal{F}_{SP}$ with SP-type round function, if P is invertible and there exists a permutation π operating on t elements such that*

$$P(x_0, \ldots, x_{t-1}) = \pi^{-1} \circ P^T \circ \pi(x_0, \ldots, x_{t-1}),$$

where $(x_0, \ldots, x_{t-1}) \in \mathbb{F}_p^t$, then there is a one-to-one correspondence between impossible differentials and zero-correlation linear hulls for the structure $\mathcal{F}_{SP}$.

Proof. As the permutation π makes P and P^T equivalent, we can transform the structure $\mathcal{F}_{SP}$ by using P^T and permutation π, which is depicted in Figure 5(c). According Corollary 2, $\mathcal{F}_{P^TS}$ is equivalent to $\mathcal{F}_{SP^T}$ with invertible P. Naturally, $\mathcal{F}_{SP}$ is equivalent to $\mathcal{F}_{SP}^{\perp}$. By using Theorem 1, we can end the proof. □

Corollary 4. *For an SPN structure $\mathcal{E}_{SP}$, if $P^TP = diag(Q_1, ..., Q_t) = Q$, where $Q_i \in \mathbb{F}_p \backslash \{0\}$, then there is a one-to-one correspondence between impossible differentials and zero-correlation linear hulls for the structure $\mathcal{E}_{SP}$.*

Proof. As $P = Q(P^{-1})^T$, for structure $\mathcal{E}_{SP}$, if substituting S by applying Q_i^{-1} before the i-th S-box of S', we have the following equivalent relation

$$\mathcal{E}_{SP} = P \circ S = (Q \circ (P^{-1})^T) \circ S = Q \circ ((P^{-1})^T \circ S') \circ Q^{-1} = Q \circ \mathcal{E}_{S(P^{-1})^T} \circ Q^{-1}.$$

Based on Definition 6, we have $\mathcal{E}_{SP}$ equivalent to its dual structure $\mathcal{E}_{S(P^{-1})^T}$. □

Corollary 5. *For a structure $\mathcal{GF}_{FP}$, if there exits a permutation π on t elements such that $F^T = \pi^{-1} \circ F \circ \pi$ and $(P^{-1})^T = \pi^{-1} \circ P \circ \pi$, then there is a one-to-one correspondence between impossible differentials and zero-correlation linear hulls for the structure $\mathcal{GF}_{FP}$.*

Proof. According to the definition of equivalence relations in [BMT13, Definition 2, Theorem 3] and Theorem 3, we can end the proof. □

When taking Corollary 3, 4, 5 and the inverse structure into consideration, we propose more refined links as follows, also depicted in Figure 6. We note that Theorem 5 works for both $\mathbb{F}_2^n$ and $\mathbb{F}_p$, it explains why some constructions have the same number of rounds in terms of the longest IDC and ZC.

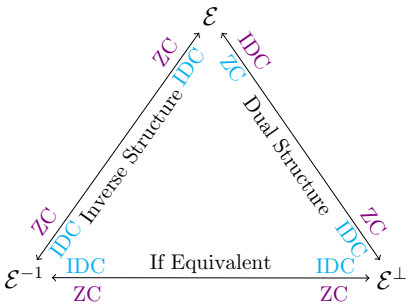


Figure 6: IDC and ZC transformations between the structure, its dual and inverse. (This figure covers the IDC and ZC part of Figure 1.)

Theorem 5. *Let $\mathcal{E} \in \{\mathcal{F}_{SP}, \mathcal{GF}_{FP}, \mathcal{E}_{SP}, \mathcal{E}_{erf}, \mathcal{E}_{crf}\}$, if its dual structure $\mathcal{E}^{\perp}$ is equivalent to the structure $\mathcal{E}$ or its inverse structure $\mathcal{E}^{-1}$, then there is a one-to-one correspondence between impossible differentials and zero-correlation linear hulls for the structure $\mathcal{E}$.*

Inspired by the link between IDC and ZC, the similar link for Prob-one DC and LC can be obtained as below.

Theorem 6. *Let $\mathcal{E} \in \{\mathcal{F}_{SP}, \mathcal{GF}_{FP}, \mathcal{E}_{SP}, \mathcal{E}_{erf}, \mathcal{E}_{crf}\}$, $a \to b$ is a Prob-one differential trail of $\mathcal{E}$ if and only if it is a Prob-one linear trail of $\mathcal{E}^{\perp}$.*

Proof. For a given structure $\mathcal{E}$, Prob-one DC (LC) means no differential (linear) active S-box in the trail for all $E \in \mathcal{E}$. Then for a given Prob-one DC of $\mathcal{E}$, it always leads to a Prob-one LC of its dual $\mathcal{E}^{\perp}$, because the input differences for all S-boxes in the cipher structure $\mathcal{E}$ are all zero and a Prob-one LC can be derived from this trail for $\mathcal{E}^{\perp}$. Vice versa. □

3.2 An Alternative Proof of Links between ZC and INT over $\mathbb{F}_p$

Recently, from a geometrical point of view of linear cryptanalysis, Beyne [Bey21] generalizes the links between zero-correlation and integral attacks, which is discovered by Bogdanov *et al.* [BLNW12] and also discussed by Sun *et al.* [SLR+15]. In this section, we explore the detailed conditions and properties of the transformation, and present alternative proofs of links between ZC and INT over $\mathbb{F}_p$. Before presenting the links, we explain the independency of input and output masks (differences) by the following definition.

Definition 8. We say that the input mask (difference) set A and output mask (difference) set B are independent, if and only if, for any $a \in A$ and any $b \in B$, $a \to b$ is a zero-correlation linear hull (impossible differential).

Lemma 3. *Let A be a subspace of $\mathbb{F}_p^t$, its orthogonal space $A^{\perp} = \{x \in \mathbb{F}_p^t | a^T \cdot x = 0, a \in A\}$. Let $F(x) : \mathbb{F}_p^t \to \mathbb{F}_p^t$ be a function over $\mathbb{F}_p^t$. For any $\lambda \in \mathbb{F}_p^t$, function $G_\lambda : A^{\perp} \mapsto \mathbb{F}_p^t$ is defined as $G_\lambda(x) = F(x + \lambda)$, then for any mask $b \in \mathbb{F}_p^t$,*

$$cor(-b^T \cdot G_\lambda(x)) = \sum_{a \in A} e^{\frac{-2\pi i}{p}(a^T \cdot \lambda)} cor(a^T \cdot x - b^T \cdot F(x)).$$

Proof. For the subspace A of $\mathbb{F}_p^t$, the equation below can be firstly deduced

$$\sum_{a \in A} e^{\frac{2\pi i}{p}(a^T \cdot x)} = \begin{cases} |A|, & \text{if } x \in A^{\perp}, \\ 0, & \text{if } x \notin A^{\perp}. \end{cases}$$

Then according to the Definition 1, it has the following

$$\begin{aligned} cor(-b^T \cdot G_\lambda(x)) &= \frac{1}{|A^\perp|} \sum_{x \in A^\perp} e^{\frac{2\pi i}{p}(-b^T \cdot G_\lambda(x))} = \frac{1}{p^t} \sum_{x \in A^\perp} (e^{\frac{2\pi i}{p}(-b^T \cdot G_\lambda(x))}|A|) \\ &= \frac{1}{p^t} \sum_{x \in A^\perp} (e^{\frac{2\pi i}{p}(-b^T \cdot G_\lambda(x))} \sum_{a \in A} e^{\frac{2\pi i}{p}(a^T \cdot x)}) \\ &= \frac{1}{p^t} \sum_{x \in \mathbb{F}_p^t} (e^{\frac{2\pi i}{p}(-b^T \cdot F(x+\lambda))} \sum_{a \in A} e^{\frac{2\pi i}{p}(a^T \cdot x)}) \end{aligned}$$

Now, let $x + \lambda = z$

$$\begin{aligned} cor(-b^T \cdot G_\lambda(x)) &= \sum_{a \in A} (\frac{1}{p^t} \sum_{z-\lambda \in \mathbb{F}_p^t} e^{\frac{2\pi i}{p}(a^T \cdot z - b^T \cdot F(z) - a^T \cdot \lambda)}) \\ &= \sum_{a \in A} e^{\frac{-2\pi i}{p}(a^T \cdot \lambda)} cor(a^T \cdot z - b^T \cdot F(z)). \end{aligned}$$

Thus, we have $cor(-b^T \cdot G_\lambda(x)) = \sum_{a \in A} e^{\frac{-2\pi i}{p}(a^T \cdot \lambda)} cor(a^T \cdot x - b^T \cdot F(x))$. □

With the input mask space A for ZC, the input space $A^\perp$ for INT and the defined function G_λ to cancel the effect of constants, the transformation from ZC to INT over $\mathbb{F}_p$ can be naturally obtained.

Theorem 7. *If there exists a subspace A of $\mathbb{F}_p^t$ and a mask $b \in \mathbb{F}_p^t \backslash \{0\}$, such that for any $a \in A$, $cor(a^T \cdot x - b^T \cdot F(x)) = 0$ where $x \in \mathbb{F}_p^t$, then for any $\lambda \in \mathbb{F}_p^t$, $b^T \cdot G_\lambda(x)$ is balanced on the subspace $A^\perp$, that is $cor(-b^T \cdot G_\lambda(x)) = 0$.*

Proof. As $cor(a^T \cdot x - b^T \cdot F(x)) = 0$ where $x \in \mathbb{F}_p^t$ for any $a \in A$, then according to Lemma 3, we can end the proof. □

Theorem 7 reveals the relation from ZC to INT and the exact form of the transformed INT. Furthermore, as required in [BLNW12], "*input and output linear masks in zero-correlation approximations are independent*", this condition over $\mathbb{F}_2^n$ later can be relaxed in [SLR+15]. However, from Lemma 3 and Theorem 7 presented above, it requires a subspace A for the input mask, that means for any $a \in A$, $a \to b$ is a zero-correlation linear hull. It can be observed that this independent condition over $\mathbb{F}_p^t$ for the input and output masks of zero-correlation linear hull cannot be removed, because the smallest nontrivial subspace of $\mathbb{F}_p^t$ has the size of p, and it has $(p-1)$ nontrivial zero-correlation linear hulls. While over $\mathbb{F}_2^n$, it only needs any one nontrivial zero-correlation linear hull $a \to b$ then $\{a, 0\}$ forms a nontrivial subspace of $\mathbb{F}_2^n$, which exhibits the gap between $\mathbb{F}_p$ and $\mathbb{F}_2^n$. In the following, we focus on the specific conditions and properties of INT that can lead to ZC. The detailed proof of Lemma 4 is provided in Appendix B, then combined with Theorem 7, Theorem 8 is obtained.

Lemma 4. *Let A be a subspace of $\mathbb{F}_p^t$, its orthogonal space is $A^\perp = \{x \in \mathbb{F}_p^t | a^T \cdot x = 0, a \in A\}$. Let $F(x) : \mathbb{F}_p^t \to \mathbb{F}_p^t$ be a function over $\mathbb{F}_p^t$. For $\lambda \in \mathbb{F}_p^t$, function $G_\lambda : A^\perp \mapsto \mathbb{F}_p^t$ is defined as $G_\lambda(x) = F(x + \lambda)$, then for any $b \in \mathbb{F}_p^t$,*

$$\frac{1}{p^t} \sum_{\lambda \in \mathbb{F}_p^t} e^{\frac{2\pi i}{p}(b^T \cdot F(\lambda))} cor(b^T \cdot G_\lambda(x)) = \sum_{a \in A} |cor(a^T \cdot x - b^T \cdot F(x))|^2.$$

Theorem 8. *Let $E(x) : \mathbb{F}_p^t \to \mathbb{F}_p^t$ be a function over $\mathbb{F}_p^t$, A be a nontrivial subspace of $\mathbb{F}_p^t$ and its orthogonal space $A^\perp = \{x \in \mathbb{F}_p^t | a^T \cdot x = 0, a \in A\}$. For any $\lambda \in \mathbb{F}_p^t$, function*

$G_\lambda : A^\perp \mapsto \mathbb{F}_p^t$ is defined as $G_\lambda(x) = E(x + \lambda)$. Then an integral distinguisher of E can lead to a zero-correlation linear hull with input masks A and nonzero output mask b, if and only if it is a balanced integral distinguisher with $b^T \cdot G_\lambda(x)$ balanced on the subspace $A^\perp$.

Proof. We consider following two parts.

- If an integral distinguisher can be transformed into a zero-correlation linear hull with input masks A, then for non-zero output mask b, we obtain

$$\sum_{a \in A} |cor(a^T \cdot x - b^T \cdot E(x))|^2 = 0.$$

  According to Theorem 7, for any $\lambda \in \mathbb{F}_p^t$, it has $cor(-b^T \cdot G_\lambda(x)) = 0$, which means that $b^T \cdot G_\lambda(x)$ is balanced on $A^\perp$ for this integral distinguisher.

- For an integral distinguisher that $b^T \cdot G_\lambda(x)$ is balanced on the subspace $A^\perp$, then according to Lemma 4, we have the following

$$\frac{1}{p^t} \sum_{\lambda \in \mathbb{F}_p^t} e^{\frac{2\pi i}{p}(b^T \cdot F(\lambda))} cor(b^T \cdot G_\lambda(x)) = 0,$$

  which leads to a zero-corelation linear hull $A \to b$.

As claimed. □

3.3 Links between IDC and INT over $\mathbb{F}_p$

According to the links presented above, now the links between IDC and INT over $\mathbb{F}_p$ can be easily established, which also has the independent conditions brought from the links of ZC and INT over $\mathbb{F}_p$. As indicated in Theorem 7, the input space $A^\perp$ for INT is the orthogonal space of the input mask space A for ZC, we do not specify the distinguishers in this subsection.

Theorem 9. *Let $\mathcal{E} \in \{\mathcal{F}_{SP}, \mathcal{GF}_{FP}, \mathcal{E}_{SP}, \mathcal{E}_{erf}, \mathcal{E}_{crf}\}$, then an impossible differential of $\mathcal{E}$ always implies the existence of an integral of $\mathcal{E}^\perp$, if its input and output differences are independent as defined in Definition 8.*

Proof. The transformation from IDC to INT can be divided into two parts: 1) from IDC to ZC (Theorem 1-4); 2) from ZC to INT (Theorem 7). □

In case $\mathcal{E}^\perp = \pi \circ \mathcal{E} \circ \pi'$ where π and π' are linear transformations, some more refined links are presented as follows.

Corollary 6. *Let $\mathcal{F}_{SP}$ be a Feistel structure with SP-type round function, and let the linear transformation be P. If P is invertible and there exists a permutation π operating on t elements such that $P(x_0, \ldots, x_{t-1}) = \pi^{-1} \circ P^T \circ \pi(x_0, \ldots, x_{t-1})$, where $(x_0, \ldots, x_{t-1}) \in \mathbb{F}_p^t$. Then for $\mathcal{F}_{SP}$, an impossible differential always implies the existence of an integral distinguisher, if its input and output differences are independent.*

Proof. Based on Corallary 3 from IDC to ZC, it has from ZC to INT by Theorem 9. □

Corollary 7. *Let $\mathcal{E}_{SP}$ be an SPN structure with the linear transformation being P. If $P^T P = diag(Q_1, ..., Q_t)$, where $Q_i \in \mathbb{F}_p \backslash \{0\}$, then for $\mathcal{E}_{SP}$, an impossible differential always implies the existence of an integral distinguisher, if its input and output differences are independent.*

Proof. Based on Corallary 4 from IDC to ZC, it has from ZC to INT by Theorem 9. □

Corollary 8. *Let $\mathcal{GF}_{FP}$ be a Generalized Feistel structure, if there exits a permutation π on t elements such that $F^T = \pi^{-1} \circ F \circ \pi$ and $(P^{-1})^T = \pi^{-1} \circ P \circ \pi$, then an impossible differential always implies the existence of an integral distinguisher, if its input and output differences are independent.*

Proof. Based on Corallary 5 from IDC to ZC, it has from ZC to INT by Theorem 9. □

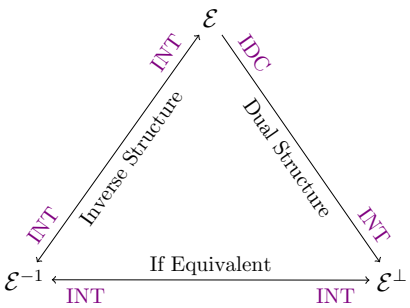


Figure 7: IDC and INT transformation between the structure, its dual and inverse. (This figure covers the IDC and INT part of Figure 1.)

Similarly, when considering all the structure, its dual and inverse structures, we have the following refined links, also as shown in Figure 7.

Theorem 10. *For a structure $\mathcal{E} \in \{\mathcal{F}_{SP}, \mathcal{GF}_{FP}, \mathcal{E}_{SP}, \mathcal{E}_{erf}, \mathcal{E}_{crf}\}$, if its dual structure $\mathcal{E}^{\perp}$ is equivalent to the structure $\mathcal{E}$ or its inverse $\mathcal{E}^{-1}$, then for $\mathcal{E}$, an impossible differential always implies the existence of an integral distinguisher, if its input and output differences are independent.*

4 Equation-based Method of Finding IDC/ZC and Applications of Links for GMiMC

In this section, as applications of the comprehensive links presented in previous section, we first utilize the equation-based method to find impossible differential and/or zero-correlation linear hull for GMiMC, then different types of improved distinguishers that are derived from the links can be achieved for all GMiMC constructions.

4.1 Impossible differential of GMiMC$_{\texttt{erf}}$ over $\mathbb{F}_p$

For GMiMC$_{\texttt{erf}}$ with number of branches t, intuitively, to have more deterministic rounds, its IDC can be divided into three parts,

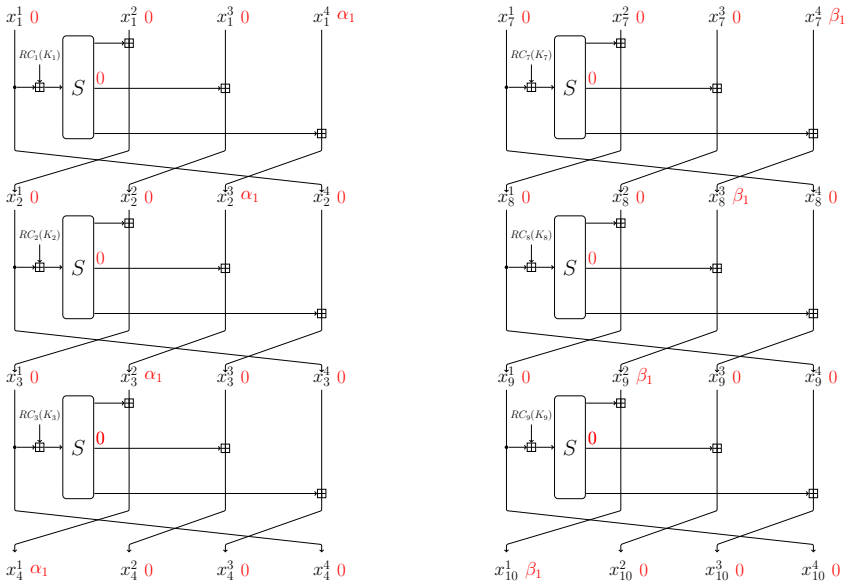
- Forward: the first $(t-1)$ rounds with probability one;
- Middle: the middle $r_1 + r_2$ $(1 \leq r_1, r_2 \leq t)$ rounds with contradictions;
- Backward: the last $(t-1)$ rounds with probability one.

Considering the first part (an example of $t = 4$ is depicted in Figure 8(a)), we denote the input difference by $\Delta_1^{forward} = (0, \cdots, 0, \alpha_1)$, then it passes $(t-1)$ rounds to the output difference $\Delta_t^{forward} = (\alpha_1, 0, \cdots, 0)$, that is,

$$\Delta_1^{forward} = (0, \cdots, 0, \alpha_1) \xrightarrow{(t-1)-round} \Delta_t^{forward} = (\alpha_1, 0, \cdots, 0).$$

Similarly, for the last part (an example of $t = 4$ is depicted in Figure 8(b)), it has $(t-1)$ free rounds backwards,

$$\nabla_t^{backward} = (0, \cdots, 0, \beta_1) \xleftarrow{(t-1)-round} \nabla_1^{backward} = (\beta_1, 0, \cdots, 0).$$



(a) Forward $(t-1)$ rounds. (b) Backward $(t-1)$ rounds.

Figure 8: First and last parts of IDC of GMiMC$_{\texttt{erf}}$ with $t = 4$.

For the middle part (an example of $t = 4$ is depicted in Figure 9), after r_1 $(1 \le r_1 \le t)$ rounds forwards, the input difference $\Delta_1^{middle} = (\alpha_1, 0, \cdots, 0)$ will lead to the output difference $\Delta_{1+r_1}^{middle}$

$$(\underbrace{\sum_{2\le i\le r_1+1}\alpha_i, \cdots, \sum_{2\le i\le r_1+1}\alpha_i}_{t-r_1}, \underbrace{\alpha_1 + \sum_{2\le i\le r_1+1}^{i\neq 2}\alpha_i, \sum_{2\le i\le r_1+1}^{i\neq 3}\alpha_i, \cdots, \sum_{2\le i\le r_1+1}^{i\neq r_1+1}\alpha_i}_{r_1}).$$

After r_2 $(1 \le r_2 \le t)$ rounds backwards, the output difference $\nabla_1^{middle} = (0, 0, \cdots, \beta_1)$ will lead to the input difference $\nabla_{1+r_2}^{middle}$

$$(\underbrace{\sum_{2\le i\le r_2+1}^{i\neq r_2+1}-\beta_i, \cdots, \sum_{2\le i\le r_2+1}^{i\neq 3}-\beta_i, \beta_1 + \sum_{2\le i\le r_2+1}^{i\neq 2}-\beta_i}_{r_2}, \underbrace{\sum_{2\le i\le r_2+1}-\beta_i, \cdots, \sum_{2\le i\le r_2+1}-\beta_i}_{t-r_2}).$$

Naturally, for a valid differential trail, the output difference $\Delta_{1+r_1}^{middle}$ and the input difference $\nabla_{1+r_2}^{middle}$ meeting in the middle part should be equal,

$$\Delta_{1+r_1}^{middle} = \nabla_{1+r_2}^{middle}. \tag{1}$$

Conversely, if we find some contradictions in the equation system (1), it will lead to an IDC with $(2t - 2 + r_1 + r_2)$ rounds for GMiMC$_{\texttt{erf}}$.

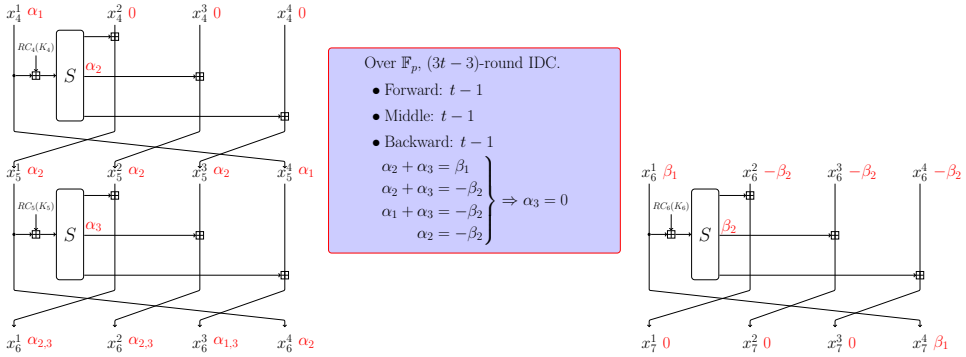


Figure 9: Middle part of IDC of $\text{GMiMC}_{\texttt{erf}}$ with $t=4$.

4.1.1 IDC of $\text{GMiMC}_{\texttt{erf}}$ with $(3t-3)$ Rounds

According to Lemma 5, we can obtain a nontrivial r-round ($2t+1 \le r \le 3t-3$) IDC of $\text{GMiMC}_{\texttt{erf}}$ with input difference $(0,\cdots,0,\alpha_1)$ and output difference $(\beta_1,0,\cdots,0)$, where $\alpha_1 \neq 0$ and $\beta_1 \neq 0$. To be more specific, due to $\alpha_1 \xrightarrow{S} \alpha_2 \xrightarrow{S} \alpha_3$, $\beta_1 \xrightarrow{S} \beta_2 \xrightarrow{S} \beta_3$ and S takes the power map $x \mapsto x^3$ as a non-linear permutation. Thus, $\alpha_3=0$ or $\beta_3=0$ will result in contradictions of $\alpha_1 \neq 0$ or $\beta_1 \neq 0$.

Lemma 5. *When $3 \le r_1+r_2 \le t-1$, the equation system (1) will lead to $\alpha_3=0$ or $\beta_3=0$.*

Proof. If we have α_3 appearing in the equation system (1), that is $2 \le r_1$ and $3 \le r_1+r_2$. Considering the rightmost (r_1+1) consecutive blocks in the output difference $\Delta^{middle}_{1+r_1}$

$$(\cdots, \sum_{2\le i\le r_1+1}\alpha_i, \underbrace{\alpha_1+\sum_{2\le i\le r_1+1}^{i\neq 2}\alpha_i, \sum_{2\le i\le r_1+1}^{i\neq 3}\alpha_i, \cdots, \sum_{2\le i\le r_1+1}^{i\neq r_1+1}\alpha_i}_{r_1}),$$

and the rightmost $(t-r_2)$ consecutive blocks in the input difference $\nabla^{middle}_{1+r_2}$

$$(\cdots, \underbrace{\sum_{2\le i\le r_2+1}-\beta_i, \cdots, \sum_{2\le i\le r_2+1}-\beta_i}_{t-r_2}).$$

If $r_1+1 \le t-r_2$, the last (r_1+1) equations in the system (1) will be

$$\begin{cases} \sum\limits_{2\le i\le r_1+1}\alpha_i = \sum\limits_{2\le i\le r_2+1}-\beta_i, \\ \alpha_1+\sum\limits_{2\le i\le r_1+1}^{i\neq 2}\alpha_i = \sum\limits_{2\le i\le r_2+1}-\beta_i, \\ \sum\limits_{2\le i\le r_1+1}^{i\neq 3}\alpha_i = \sum\limits_{2\le i\le r_2+1}-\beta_i, \\ \cdots \\ \sum\limits_{2\le i\le r_1+1}^{i\neq r_1+1}\alpha_i = \sum\limits_{2\le i\le r_2+1}-\beta_i. \end{cases} \tag{2}$$

Then, we can easily deduce $\alpha_3 = 0$ from the first and third equations of (2),

$$\sum_{2\le i\le r_1+1} \alpha_i = \sum_{2\le i\le r_2+1} -\beta_i = \sum_{2\le i\le r_1+1}^{i\neq 3} \alpha_i.$$

Similarly, if we have β_3 in the system (1), we can also deduce $\beta_3 = 0$. □

4.1.2 IDC of GMiMC$_{\texttt{erf}}$ with $(3t-1)$ Rounds

Similarly, according to Lemma 6, we can extend two more rounds for IDC of GMiMC$_{\texttt{erf}}$.

Lemma 6. *When $3 \le r_1 + r_2 \le t+1$, $\alpha_1 = \beta_1$ and $t \not\equiv 1 \bmod p$, the equation system (1) will lead to $\alpha_3 = 0$ or $\beta_3 = 0$.*

Proof. We consider the following three cases, where the conditions that $\alpha_1 = \beta_1$ and $t \not\equiv 1 \bmod p$ are used in the last two cases.

- **Case 1.** When $3 \le r_1 + r_2 \le t-1$, it has $\alpha_3 = 0$ or $\beta_3 = 0$ by using Lemma 5.
- **Case 2.** When $r_1 + r_2 = t$, it has the following equations,

$$\begin{cases} \sum_{2\le i\le r_1+1} \alpha_i = \sum_{2\le i\le r_2+1}^{i\neq r_2+1} -\beta_i, \\ \qquad \cdots, \\ \sum_{2\le i\le r_1+1} \alpha_i = \sum_{2\le i\le r_2+1}^{i\neq 3} -\beta_i, \\ \sum_{2\le i\le r_1+1} \alpha_i = \beta_1 + \sum_{2\le i\le r_2+1}^{i\neq 2} -\beta_i, \\ \alpha_1 + \sum_{2\le i\le r_1+1}^{i\neq 2} \alpha_i = \sum_{2\le i\le r_2+1} -\beta_i, \\ \sum_{2\le i\le r_1+1}^{i\neq 3} \alpha_i = \sum_{2\le i\le r_2+1} -\beta_i, \\ \qquad \cdots, \\ \sum_{2\le i\le r_1+1}^{i\neq r_1+1} \alpha_i = \sum_{2\le i\le r_2+1} -\beta_i. \end{cases} \tag{3}$$

If $r_2 = 1$, then $r_1 = t-1$, based on $\alpha_1 = \beta_1$ and equation system (3),

$$\begin{cases} (t-1)\sum_{2\le i\le r_1+1} \alpha_i = -(t-1)\beta_2, \\ \alpha_2 = -\beta_2, \\ \alpha_3 = \cdots = \alpha_{r_1+1}. \end{cases}$$

If $t \not\equiv 1 \bmod p$, then there is $\alpha_3 = 0$. In the same way, if $r_1 = 1$, then $r_2 = t-1$, it will lead to $\beta_3 = 0$. Now for $2 \le r_1, r_2$, it has the following,

$$\begin{cases} (t-1)\sum_{2\le i\le r_1+1} \alpha_i + (t-1)\sum_{2\le i\le r_2+1} \beta_i = 0, \\ \alpha_3 = \cdots = \alpha_{r_1+1}, \\ \beta_3 = \cdots = \beta_{r_2+1}, \\ \alpha_3 = \beta_3. \end{cases}$$

Still, if $t \not\equiv 1 \bmod p$, $\alpha_3 = \beta_3 = 0$ can be easily deduced.

- **Case 3.** When $r_1 + r_2 = t + 1$, the blocks in the output difference $\Delta^{middle}_{1+r_1}$ can be divided into three parts as below,

$$(\underbrace{\sum_{2\le i\le r_1+1}\alpha_i,\cdots,\sum_{2\le i\le r_1+1}\alpha_i}_{t-r_1},\alpha_1+\sum_{2\le i\le r_1+1}^{i\neq 2}\alpha_i,\underbrace{\sum_{2\le i\le r_1+1}^{i\neq 3}\alpha_i,\cdots,\sum_{2\le i\le r_1+1}^{i\neq r_1+1}\alpha_i}_{r_1-1}),$$

and also for the input difference $\nabla^{middle}_{1+r_2}$,

$$(\underbrace{\sum_{2\le i\le r_2+1}^{i\neq r_2+1}-\beta_i,\cdots,\sum_{2\le i\le r_2+1}^{i\neq 3}-\beta_i}_{r_2-1},\beta_1+\sum_{2\le i\le r_2+1}^{i\neq 2}-\beta_i,\underbrace{\sum_{2\le i\le r_2+1}-\beta_i,\cdots,\sum_{2\le i\le r_2+1}-\beta_i}_{t-r_2}).$$

As $r_1 - 1 = t - r_2$ and $r_2 - 1 = t - r_1$, we can obtain the following equations,

$$\begin{cases}\alpha_1+\sum\limits_{2\le i\le r_1+1}^{i\neq 2}\alpha_i=\beta_1+\sum\limits_{2\le i\le r_2+1}^{i\neq 2}-\beta_i,\\ \sum\limits_{2\le i\le r_2+1}^{i\neq r_2+1}-\beta_i=\sum\limits_{2\le i\le r_1+1}\alpha_i,\\ \cdots,\\ \sum\limits_{2\le i\le r_2+1}^{i\neq 3}-\beta_i=\sum\limits_{2\le i\le r_1+1}\alpha_i,\\ \sum\limits_{2\le i\le r_1+1}^{i\neq 3}\alpha_i=\sum\limits_{2\le i\le r_2+1}-\beta_i,\\ \cdots,\\ \sum\limits_{2\le i\le r_1+1}^{i\neq r_1+1}\alpha_i=\sum\limits_{2\le i\le r_2+1}-\beta_i.\end{cases}\tag{4}$$

If $r_2 = 1$ and $r_1 = t$, then based on $\alpha_1 = \beta_1$ and equations (4), we have,

$$\begin{cases}\sum\limits_{3\le i\le r_1+1}\alpha_i=0,\\ \alpha_3=\cdots=\alpha_{r_1+1}.\end{cases}\tag{5}$$

If $t \not\equiv 1 \bmod p$, $\alpha_3 = 0$ can be deduced. In the same way, if $r_1 = 1$, then $r_2 = t$, it can be deduced that $\beta_3 = 0$. Now for $2 \le r_1, r_2$, we have the following equations,

$$\begin{cases}\sum\limits_{3\le i\le r_1+1}\alpha_i+\sum\limits_{3\le i\le r_2+1}\beta_i=0,\\ \alpha_3=\cdots=\alpha_{r_1+1},\\ \beta_3=\cdots=\beta_{r_2+1},\\ \alpha_3=\beta_3.\end{cases}$$

Still, if $t \not\equiv 1 \bmod p$, $\alpha_3 = \beta_3 = 0$ can be easily deduced.

Considering all three cases above, it has $\alpha_3 = 0$ or $\beta_3 = 0$. □

An example of this difference propagation and equation system is depicted in Figure 10.

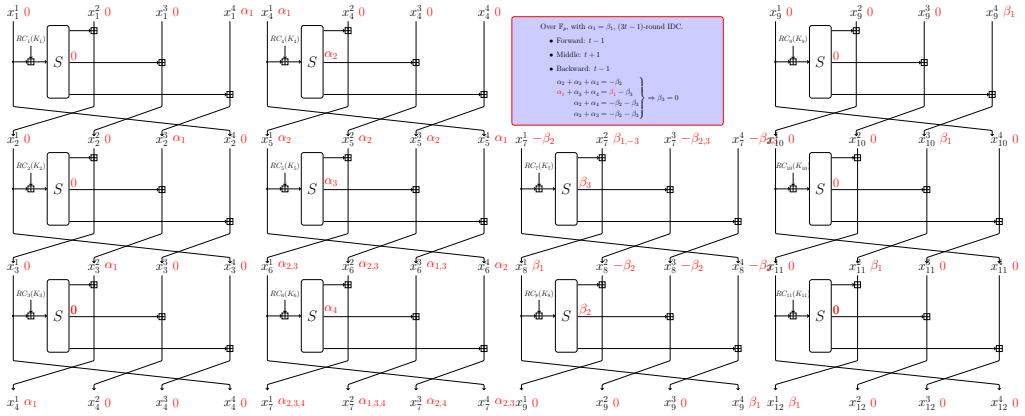


Figure 10: $(3t-1)$ rounds IDC example of GMiMC$_{\texttt{erf}}$ with number of branch $t=4$.

4.1.3 IDC of GMiMC$_{\texttt{erf}}$ with an *arbitrary number of rounds*

Different from $(3t-1)$ and $(3t-3)$ rounds IDC of GMiMC$_{\texttt{erf}}$ presented above, we now present a special case over $\mathbb{F}_p$ provided in Lemma 7, which can lead to an arbitrary number of rounds IDC of GMiMC$_{\texttt{erf}}$ and only works over $\mathbb{F}_p$. Although it has the limited condition that $t \equiv 1 \bmod p$, it shows the difference between $\mathbb{F}_2^n$ and $\mathbb{F}_p$. To be specific, this can be attributed to that $\alpha_1 = -\beta_1$ is equivalent to $\alpha_1 = \beta_1$ over $\mathbb{F}_2^n$. However, this will be different over $\mathbb{F}_p$, $\alpha_1 = -\beta_1$ combined with $\alpha_1 = \beta_1$ will lead to $\alpha_1 = \beta_1 = 0$, which can be used to construct an arbitrary number of rounds IDC of GMiMC$_{\texttt{erf}}$ with input difference $(0, \cdots, 0, \alpha_1)$ and output difference $(\beta_1, 0, \cdots, 0)$, where $\alpha_1 = -\beta_1 \neq 0$.

Lemma 7. *When $\alpha_1 = -\beta_1$ and $t \equiv 1 \bmod p$, for any r_1, r_2 $(1 \leq r_1, r_2)$, the equation system (1) will lead to $\alpha_1 = \beta_1 = 0$.*

Proof. For any r_1, r_2 $(1 \leq r_1, r_2)$ of the middle part mentioned above, with the input difference $\Delta_1^{middle} = (\alpha_1, 0, \cdots, 0)$ and the output difference $\nabla_1^{middle} = (0, 0, \cdots, \beta_1)$, we then generalize the representations of the difference $\Delta_{1+r_1}^{middle}$

$$\Big(\underbrace{\sum_{2\leq i\leq r_1+1}^{i\neq r_1+2 \bmod t}\alpha_i, \cdots, \sum_{2\leq i\leq r_1+1}^{i\neq t+1 \bmod t}\alpha_i}_{(t-r_1) \bmod t}, \alpha_1 + \underbrace{\sum_{2\leq i\leq r_1+1}^{i\neq 2 \bmod t}\alpha_i, \sum_{2\leq i\leq r_1+1}^{i\neq 3 \bmod t}\alpha_i, \cdots, \sum_{2\leq i\leq r_1+1}^{i\neq r_1+1 \bmod t}\alpha_i}_{(r_1-1) \bmod t}\Big),$$

and the difference $\nabla_{1+r_2}^{middle}$

$$\Big(\underbrace{\sum_{2\leq i\leq r_2+1}^{i\neq r_2+1 \bmod t}-\beta_i, \cdots, \sum_{2\leq i\leq r_2+1}^{i\neq 3 \bmod t}-\beta_i}_{(r_2-1) \bmod t}, \beta_1 + \underbrace{\sum_{2\leq i\leq r_2+1}^{i\neq 2 \bmod t}-\beta_i, \sum_{2\leq i\leq r_2+1}^{i\neq t+1 \bmod t}-\beta_i, \cdots, \sum_{2\leq i\leq r_2+1}^{i\neq r_2+2 \bmod t}-\beta_i}_{(t-r_2) \bmod t}\Big).$$

Then for a valid differential trail, the differences $\Delta_{1+r_1}^{middle}$ and $\nabla_{1+r_2}^{middle}$ should meet in the middle, that is $\Delta_{1+r_1}^{middle} = \nabla_{1+r_2}^{middle}$. When sum all terms in each block of $\Delta_{1+r_1}^{middle}$, it has

$$sum(\Delta_{1+r_1}^{middle}) = \alpha_1 + (t-1)\cdot\alpha_2 + (t-1)\cdot\alpha_3 + \cdots (t-1)\cdot\alpha_{r_1+1}. \tag{6}$$

Similarly, we have the sum for $\nabla_{1+r_2}^{middle}$,

$$sum(\nabla_{1+r_2}^{middle}) = \beta_1 - (t-1)\cdot\beta_2 - (t-1)\cdot\beta_3 - \cdots (t-1)\cdot\beta_{r_2+1}. \tag{7}$$

Naturally, we obtain an equation over $\mathbb{F}_p$ from equations (6) and (7) as below

$$\alpha_1 + (t-1)\cdot\alpha_2 + \cdots (t-1)\cdot\alpha_{r_1+1} \equiv \beta_1 - (t-1)\cdot\beta_2 - \cdots (t-1)\cdot\beta_{r_2+1} \bmod p.$$

As $t \equiv 1 \bmod p$, it has $\alpha_1 = \beta_1$, and combined with $\alpha_1 = -\beta_1$, we obtain $\alpha_1 = \beta_1 = 0$. □

4.1.4 Transformation from IDC to ZC and INT of GMiMC$_{\texttt{crf}}$

With these three IDCs of GMiMC$_{\texttt{erf}}$ presented above, now by using the link proposed in Theorem 4, we can directy obtain the corresponding ZCs of GMiMC$_{\texttt{crf}}$ over $\mathbb{F}_p$ as below

- $(3t-3)$-round: $(0,\cdots,0,a_1) \nrightarrow (b_1,0,\cdots,0)$, where $a_1 \neq 0$ and $b_1 \neq 0$.
- $(3t-1)$-round: $(0,\cdots,0,a_1) \nrightarrow (b_1,0,\cdots,0)$, where $a_1 = b_1 \neq 0$ and $t \not\equiv 1 \bmod p$.
- Arbitrary number of rounds: $(0,\cdots,0,a_1) \nrightarrow (b_1,0,\cdots,0)$, where $a_1 = -b_1 \neq 0$ and $t \equiv 1 \bmod p$.

Then, by using the link given in Theorem 9, a $(3t-3)$-round INT of GMiMC$_{\texttt{crf}}$ over $\mathbb{F}_p$ can be obtained. Let $V = \{(0,\cdots,0,x) | x \in \mathbb{F}_p\}$, if the input space is $V^{\perp}$, then the output is balanced on $(b_1,0,\cdots,0)\cdot$GMiMC$_{\texttt{crf}}$ where $b_1 \neq 0$.

4.2 Zero-correlation Linear Hull of GMiMC$_{\texttt{erf}}$ over $\mathbb{F}_p$

Linear Trail with Probability One. Before presenting the ZC of GMiMC$_{\texttt{erf}}$ over $\mathbb{F}_p$, we first discuss the possible free rounds of its linear trail. As shown in Figure 11, the input masks are denoted by the same element $a_1 \in \mathbb{F}_p \backslash \{0\}$. Then, for one free round passing, the input and output masks of a non-linear permutation S will be both zero. Based on the propagation rules of linear mask introduced in Section 2.1, a free $(t-1)$-round linear trail of GMiMC$_{\texttt{erf}}$ can be obtained with input mask $(a_1,\cdots,a_1,(2-t)a_1)$ and output mask $((2-t)a_1,a_1,\cdots,a_1)$. In fact, this is same as the linear relation proposed in [BCD+20], which can be also interpreted from the view of linear mask propagation. Now, similar to the IDC of GMiMC$_{\texttt{erf}}$, if having the limited condition on the number of branch, that is $t \equiv 1 \bmod p$, then the input and output mask of $(t-1)$-round linear trail are both $(a_1,\cdots,a_1)$, which is iterative and will lead to a linear trail with an arbitrary number of rounds. By using Theorem 6, these two linear trails with probability one can be also transformed into Prob-one differential trails of GMiMC$_{\texttt{crf}}$.

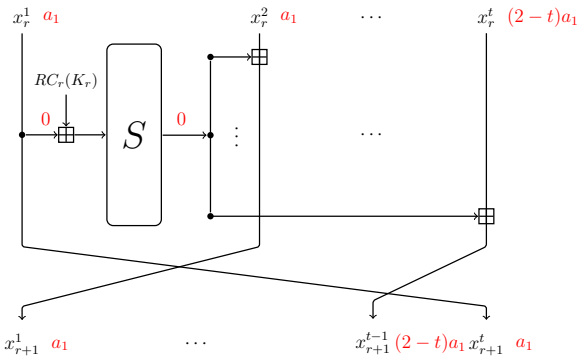


Figure 11: One free round of linear trail of GMiMC$_{\texttt{erf}}$ with t branches.

Now we briefly explain three parts of ZC of GMiMC$_{\texttt{erf}}$ as below, which is similar to that of IDC.

- Forward: the first $(t-1)$ rounds with linear probability one;

- Middle: the middle $r_1 + r_2$ $(1 \leq r_1, r_2 \leq t)$ rounds with contradictions on some variables;
- Backward: the last $(t-1)$ rounds with linear probability one.

The input mask $\Gamma_1^{forward} = (a_1, a_1, \cdots, a_1, (2-t)a_1)$ can pass $(t-1)$ free rounds to the output mask $\Gamma_t^{forward} = ((2-t)a_1, a_1, a_1, \cdots, a_1)$, that is

$$\Gamma_1^{forward} \xrightarrow{(t-1)-round} \Gamma_t^{forward}.$$

Similarly, we have free $(t-1)$ free rounds backwards

$$\Lambda_t^{backward} \xleftarrow{(t-1)-round} \Lambda_1^{backward},$$

where the input mask $\Lambda_1^{backward} = ((2-t)\cdot b_1, b_1, \cdots, b_1)$ can pass $(t-1)$ free rounds to the output mask $\Lambda_t^{backward} = (b_1, \cdots, b_1, (2-t)\cdot b_1)$.

For the middle part, after r_1 $(1 \leq r_1 \leq t)$ rounds forwards, the input mask $\Gamma_1^{middle} = ((2-t)\cdot a_1, a_1, a_1, \cdots, a_1)$ will lead to the output mask

$$\Gamma_{1+r_1}^{middle} = (\underbrace{a_1, \cdots, a_1}_{t-r_1}, \underbrace{(2-t)\cdot a_1 - a_2, a_1 - a_3, \cdots, a_1 - a_{r_1+1}}_{r_1}).$$

After r_2 $(1 \leq r_2 \leq t)$ rounds backwards, the output mask $\Lambda_1^{middle} = (b_1, \cdots, b_1, (2-t)\cdot b_1)$ will lead to the input mask

$$\Lambda_{1+r_2}^{middle} = (\underbrace{b_1 + b_{r_2+1}, \cdots, b_1 + b_3, (2-t)\cdot b_1 + b_2}_{r_2}, \underbrace{b_1, \cdots, b_1}_{t-r_2}).$$

Naturally, for a valid linear trail, the output mask $\Gamma_{1+r_1}^{middle}$ and the input mask $\Lambda_{1+r_2}^{middle}$ meeting in the middle part should be equal

$$\Gamma_{1+r_1}^{middle} = \Lambda_{1+r_2}^{middle}. \tag{8}$$

Conversely, if we find some contradictions in the equation system (8), it will lead to a ZC with $(2t - 2 + r_1 + r_2)$ rounds for GMiMC$_{\texttt{erf}}$.

4.2.1 ZC of GMiMC$_{\texttt{erf}}$ with an *arbitrary number of rounds*

With the probability one linear trail given above, whose input mask Γ_{in} and output mask Γ_{out} are both $(a_1, \cdots, a_1)$, we can obtain two kinds of ZC of GMiMC$_{\texttt{erf}}$ with an arbitrary number of rounds, nevertheless only for the very limited case $t \equiv 1 \bmod p$. One with input mask Γ_{in} and output mask Γ'_{out} where $\Gamma'_{out} \neq \{\Gamma_{out}, 0\}$, another one with output mask Γ_{out} and input mask Γ'_{in} where $\Gamma'_{in} \neq \{\Gamma_{in}, 0\}$.

4.2.2 ZC of GMiMC$_{\texttt{erf}}$ with $(3t-3)$ Rounds

According to Lemma 8, we can obtain a ZC with r $(2t+1 \leq r \leq 3t-3)$ rounds for GMiMC$_{\texttt{erf}}$, its input mask is $(a_1, a_1, \cdots, a_1, (2-t)a_1)$ and output mask is $((2-t)b_1, b_1, b_1, \cdots, b_1)$, where $a_1 \neq 0$ and $b_1 \neq 0$. The proof of this ZC for GMiMC$_{\texttt{erf}}$ is similar to that of IDC.

Lemma 8. *When $3 \leq r_1 + r_2 \leq t-1$, the equation system (8) will lead to $a_3 = 0$ or $b_3 = 0$.*

Proof. If it has a_3 appearing in the equation system (8), that is $2 \leq r_1$ and $3 \leq r_1 + r_2$. Considering the rightmost $(r_1 + 1)$ consecutive blocks in the output mask $\Gamma^{middle}_{1+r_1}$,

$$(\cdots, a_1, (t-2)\cdot a_1 - a_2, \underbrace{a_1 - a_3, \cdots, a_1 - a_{r_1+1}}_{r_1-1}),$$

and the rightmost $(t - r_2)$ consecutive blocks in the input mask $\Lambda^{middle}_{1+r_2}$,

$$(\cdots, \underbrace{b_1, \cdots, b_1}_{t-r_2}).$$

If $r_1 + 1 \leq t - r_2$, that is $r_1 + r_2 \leq t - 1$, last $(r_1 + 1)$ equations in the system (8) will be

$$\begin{cases} a_1 = b_1, \\ (t-2)\cdot a_1 - a_2 = b_1, \\ a_1 - a_3 = b_1, \\ \cdots \\ a_1 - a_{r_1+1} = b_1. \end{cases} \tag{9}$$

Then, it has $a_3 = 0$ deduced from the first and third equations of (9),

$$a_1 - b_1 = a_3 = 0.$$

Similarly, if it has b_3 in the system (8), we can also deduce that $b_3 = 0$. □

An example of this linear mask propagation and equation system is depicted in Figure 12.

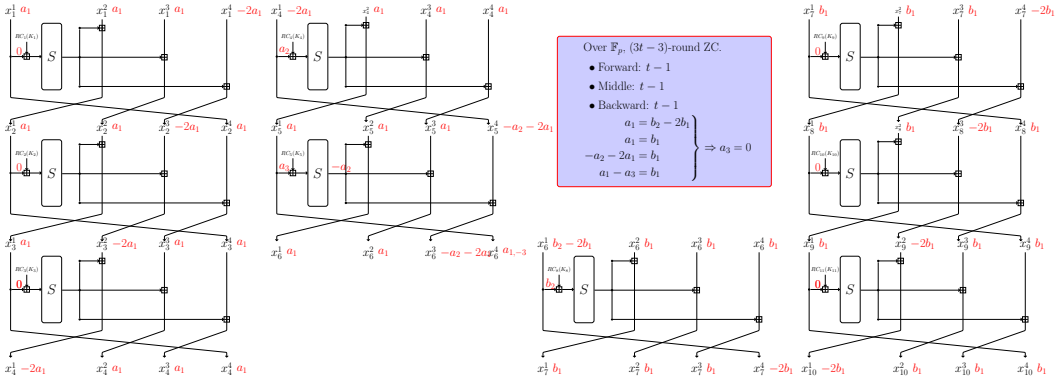


Figure 12: $(3t-3)$ rounds ZC example of GMiMC$_{\texttt{erf}}$ with branch $t = 4$.

4.2.3 ZC of GMiMC$_{\texttt{erf}}$ with $(3t-1)$ Rounds

Similarly, according to Lemma 9, we can extend two more rounds for ZC of GMiMC$_{\texttt{erf}}$.

Lemma 9. *When $3 \leq r_1 + r_2 \leq t+1$, $a_1 = b_1$, the equation system (8) will lead to $a_3 = 0$ or $b_3 = 0$.*

Proof. We consider the following three cases, where the condition $a_1 = b_1$ is used in the last two cases.

- **Case 1.** When $3 \leq r_1 + r_2 \leq t - 1$, we can easily obtain $a_3 = 0$ or $b_3 = 0$ by using Lemma 8.

- **Case 2.** When $r_1 + r_2 = t + 1$, the blocks in the output mask $\Gamma^{middle}_{1+r_1}$ can be divided into three parts as below

$$(\underbrace{a_1, \cdots, a_1}_{t-r_1}, (2-t)\cdot a_1 - a_2, \underbrace{a_1 - a_3, \cdots, a_1 - a_{r_1+1}}_{r_1-1}),$$

and similar for the input mask $\Lambda^{middle}_{1+r_2}$,

$$(\underbrace{b_1 + b_{r_2+1}, \cdots, b_1 + b_3}_{r_2-1}, (2-t)\cdot b_1 + b_2, \underbrace{b_1, \cdots, b_1}_{t-r_2}).$$

As $r_1 - 1 = t - r_2$ and $r_2 - 1 = t - r_1$, it has the following equations,

$$\begin{cases} a_1 = b_1 + b_{r_2+1}, \\ \cdots, \\ a_1 = b_1 + b_3, \\ (2-t)\cdot a_1 - a_2 = (2-t)\cdot b_1 + b_2, \\ b_1 = a_1 - a_3, \\ \cdots, \\ b_1 = a_1 - a_{r_1+1}. \end{cases} \tag{10}$$

If $r_2 = 1$, then $r_1 = t$, based on $a_1 = b_1$ and equation system (10), it has

$$a_3 = \cdots = a_{r_1+1} = a_1 - b_1 = 0.$$

Reversely, if $r_1 = 1$ and $r_2 = t$, it has $b_3 = 0$.

Now for $2 \le r_1, r_2$, we have the following equations,

$$\begin{cases} a_3 = \cdots = a_{r_1+1} = a_1 - b_1 = 0, \\ b_3 = \cdots = b_{r_2+1} = a_1 - b_1 = 0. \end{cases}$$

Still, $a_3 = b_3 = 0$ can be easily deduced.

- **Case 3.** When $r_1 + r_2 = t$, that is $r_1 = t - r_2$ and $r_2 = t - r_1$, we have the following equations,

$$\begin{cases} a_1 = b_1 + b_{r_2+1}, \\ \cdots, \\ a_1 = b_1 + b_3, \\ a_1 = (2-t)\cdot b_1 + b_2, \\ (2-t)\cdot a_1 - a_2 = b_1, \\ a_1 - a_3 = b_1, \\ \cdots, \\ a_1 - a_{r_1+1} = b_1. \end{cases} \tag{11}$$

If $r_2 = 1$, then $r_1 = t - 1$, based on $a_1 = b_1$ and equation system (11),

$$\begin{cases} a_2 + b_2 = 0, \\ a_3 = \cdots = a_{r_1+1} = a_1 - b_1 = 0. \end{cases}$$

Thus, $a_3 = 0$ can be deduced. In the same way, if $r_1 = 1$ and $r_2 = t - 1$, it has $b_3 = 0$. Now for $2 \le r_1, r_2$, we have the following equations,

$$\begin{cases} a_2 + b_2 = 0, \\ a_3 = \cdots = a_{r_1+1} = a_1 - b_1 = 0, \\ b_3 = \cdots = b_{r_2+1} = a_1 - b_1 = 0. \end{cases}$$

Still, $a_3 = b_3 = 0$ can be easily deduced.

Considering three cases above, it has $a_3 = 0$ or $b_3 = 0$. □

An example of this linear mask propagation and equation system is depicted in Figure 13.

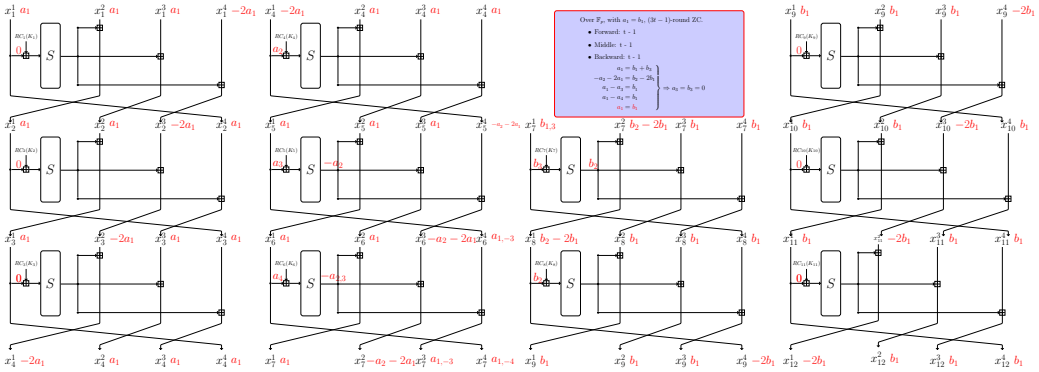


Figure 13: $(3t-1)$ rounds ZC example of GMiMC$_{\texttt{erf}}$ with branch $t = 4$.

4.2.4 Transformation from ZC to IDC of GMiMC$_{\texttt{crf}}$

With all these three kinds of ZC of GMiMC$_{\texttt{erf}}$ presented above, by using the link proposed in Theorem 4, we can obtain the corresponding IDCs of GMiMC$_{\texttt{crf}}$ as below.

- $(3t-3)$-round: with input difference $(\alpha_1, \cdots, \alpha_1, (2-t)\alpha_1)$ and output difference $((2-t)\beta_1, \beta_1, \cdots, \beta_1)$, where $\alpha_1 \neq 0$ and $\beta_1 \neq 0$.
- $(3t-1)$-round: with input difference $(\alpha_1, \cdots, \alpha_1, (2-t)\alpha_1)$ and output difference $((2-t)\beta_1, \beta_1, \cdots, \beta_1)$, where $\alpha_1 = \beta_1 \neq 0$.
- *Arbitrary number of rounds*: the input (output) difference is $(\alpha_1, \cdots, \alpha_1)$, and the non-zero output (input) difference is not equal to the input (output) difference, where $\alpha_1 \neq 0$ and $t \equiv 1 \bmod p$.

4.2.5 Transformation from ZC to INT of GMiMC$_{\texttt{erf}}$

Similarly, by using the link given in Theorem 7, we have the corresponding INTs of GMiMC$_{\texttt{erf}}$ as below.

- $(3t-3)$-round: Let $V = \{(x, x, \cdots, x, (2-t)x) | x \in \mathbb{F}_p\}$, if the input space is $V^{\perp}$, then the output is balanced on $((2-t)b_1, b_1, b_1, \cdots, b_1)\cdot$ GMiMC$_{\texttt{erf}}$ where $b_1 \neq 0$.
- *Arbitrary number of rounds*: Let $V = \{(0, \cdots, 0, x, 0, \cdots, 0) | x \in \mathbb{F}_p\}$, if the input space is V, then the output is balanced on $(b_1, \cdots, b_1)\cdot$GMiMC$_{\texttt{erf}}$ where $b_1 \neq 0$.
- *Arbitrary number of rounds*: Let $V = \{(x, \cdots, x) | x \in \mathbb{F}_p\}$, if the input space is $V^{\perp}$, any output block of GMiMC$_{\texttt{erf}}$ is balanced.

Remark: It should be noted that these probability one differential or linear trails with an *arbitrary number of rounds* presented above are *trivial*[5] for unbalanced Feistel networks, if the branch t is chosen improperly. However, in this paper, by using the equation-based methods and our proposed links, we show that bad choices of t will also lead to *nontrivial* IDC/ZC/INT with an *arbitrary number of rounds* for GMiMC. There are also some

[5]The summation of differences/masks for part of these branches can be cancelled to zero, which is similar to that over $\mathbb{F}_2^n$.

potential instantiations, for example the concrete instances with 256-bit block size and key size, GMiMC$_{\texttt{erf}}$-($p = 5, t = 86, r = 261$) and GMiMC$_{\texttt{erf}}$-($p = 17, t = 52, r = 160$)) of GMiMC$_{\texttt{erf}}$ provided in [AGP+19b, Table 6 and 7], which aim to achieve smaller signature size when intended to be deployed in post-quantum signatures with low-data scenario [CDG+17] or even full-data scenario [BEF19], and the odd number of branch has been avoided for instances over $\mathbb{F}_2^n$. We stress that considering the limited access to data, these statistical distinguishers are more suitable for the full-data setting, e.g., collision-resistant hash function [BEF19]. Nevertheless we hope these presented distinguishers could provide a guidance for future designs when considering related constructions.

4.3 Equation-based Method for GMiMC$_{\texttt{Nyb}}$

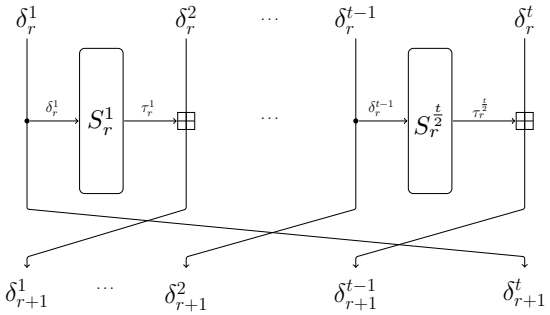


Figure 14: Differential of GMiMC$_{\texttt{Nyb}}$'s r-th round function with t branches.

For the balanced Feistel construction GMiMC$_{\texttt{Nyb}}$, we mainly focus on the underlying equivalent relations of its different structures. Specifically, we are dedicated to obtaining the one-to-one correspondence of IDC/ZC/INT for GMiMC$_{\texttt{Nyb}}$ in the following, which has already been mentioned as an example when explaining Figure 1.

One round differential propagation of GMiMC$_{\texttt{Nyb}}$ is depicted in Figure 14. The IDC of GMiMC$_{\texttt{Nyb}}$ consists of forward $r_1 (1 \leq r_1 \leq t+2)$ rounds and backward $r_2 (1 \leq r_2 \leq t+2)$ rounds. The input difference $\Delta_1^{forward} = (0, \cdots, 0, \alpha_1)$ passes r_1 rounds to the output difference $\Delta_{1+r_1}^{forward}$

$$(*, \cdots, *, \underbrace{\alpha_1 + \sum_{r+=2}^{1 \leq r \leq r_1} \tau_r^{(\frac{t}{2} - \frac{r-1}{2}) \bmod \frac{t}{2}}, *, \cdots, *}_{(r_1+1) \bmod t}),$$

where τ_r^j denotes the output difference of the j-th S-box of the r-th round. Similarly, it has the output difference $\nabla_1^{backward} = (0, \cdots, 0, \beta_1, 0)$ passing r_2 rounds backwards to $\nabla_{r_2+1}^{backward}$ as below

$$(*, \cdots, *, \underbrace{\beta_1 - \sum_{r+=2}^{1 \leq r \leq r_2} \tau_r^{'(\frac{r-1}{2}) \bmod \frac{t}{2}}, *, \cdots, *}_{(t-r_2+2) \bmod t}),$$

where $\tau_r^{'j}$ denotes the output difference of the j-th S-box of the r-th round.

Besides, if $r_2 < t$, another choice for the output difference $\nabla_1^{'backward} = (\beta_1, 0, \cdots, 0)$

will lead to $\nabla'^{backward}_{r_2+1}$ as below

$$(*, \cdots, *, 0, \underbrace{\beta_1 - \sum_{r+=2}^{1\le r\le r_2} \tau_r'^{(\frac{r+1}{2}) \bmod \frac{t}{2}}}_{(t-r_2) \bmod t}, *, \cdots, *).$$

For a valid differential trail, the output difference and the input difference meeting in the middle should be equal, thus it has the following

$$\Delta^{forward}_{1+r_1} = \nabla'^{backward}_{r_2+1}, \tag{12}$$

$$\Delta^{forward}_{1+r_1} = \nabla^{backward}_{r_2+1}. \tag{13}$$

Naturally, if we find some contradictions in the above equation systems, it will lead to an IDC with $(r_1 + r_2)$ rounds for GMiMC$_{\texttt{Nyb}}$.

4.3.1 IDC of GMiMC$_{\texttt{Nyb}}$ with $(2t - 1)$ Rounds

For any nontrivial differential trail with input difference $(0, 0, \cdots, 0, \alpha_1)$ and output difference $(\beta_1, 0, \cdots, 0)$, where $\alpha_1 \neq 0$ and $\beta_1 \neq 0$, from Lemma 10, we can obtain an IDC with $(2t - 1)$ rounds for GMiMC$_{\texttt{Nyb}}$, because all S-boxes adopted in GMiMC$_{\texttt{Nyb}}$ are permutations, and $\tau_t^{\frac{t}{2}} = 0$ or $\tau_t'^{\frac{t}{2}} = 0$ will result in contradictions of $\alpha_1 \neq 0$ or $\beta_1 \neq 0$.

Lemma 10. *When $r_1 + r_2 = 2t - 1$ with $1 \le r_1, r_2 \le t + 1$, the equation system (12) will lead to $\tau_t^{\frac{t}{2}} = 0$ or $\tau_t'^{\frac{t}{2}} = 0$.*

Proof. We may as well let $r_1 > r_2$. With the input difference $\Delta_1^{forward} = (0, \cdots, 0, \alpha_1)$, it will activate the S-box sequentially as below

$$\alpha_1 \xrightarrow{S_2^{\frac{t}{2}}} \tau_2^{\frac{t}{2}} \xrightarrow{S_3^{\frac{t}{2}}} \cdots \xrightarrow{S_t^{\frac{t}{2}}} \tau_t^{\frac{t}{2}}.$$

Then after t rounds, we have $\Delta^{forward}_{1+t} = (*, \cdots, *, \tau_t^{\frac{t}{2}}, \alpha_1)$. For the following one round, it has

$$\Delta^{forward}_{1+t+1} = (*, \cdots, *, \tau_t^{\frac{t}{2}}, \alpha_1 + \tau_{t+1}^{\frac{t}{2}}, *).$$

As for the output difference $\nabla'^{backward}_1 = (\beta_1, 0, \cdots, 0, 0)$, thus if $r_1 + 1 \equiv t - r_2 \bmod t$, that is $r_1 + r_2 \equiv -1 \bmod t$, then the block difference $\tau_t^{\frac{t}{2}}$ and 0 will coincide, which means $\tau_t^{\frac{t}{2}} = 0$. Similarly, if $r_1 < r_2$, it will lead to $\tau_t'^{\frac{t}{2}} = 0$. □

4.3.2 IDC of GMiMC$_{\texttt{Nyb}}$ with $(2t + 1)$ Rounds

For any nontrivial differential trail with input difference $(0, 0, \cdots, 0, \alpha_1)$ and output difference $(0, 0, \cdots, \beta_1, 0)$, where $\alpha_1 \neq 0$, $\beta_1 \neq 0$ and $\alpha_1 = \beta_1$, similarly from Lemma 11, an IDC with $(2t + 1)$ rounds for GMiMC$_{\texttt{Nyb}}$ can be obtained.

Lemma 11. *When $r_1 + r_2 = 2t + 1$ and $\alpha_1 = \beta_1$ with $1 \le r_1, r_2 \le t + 2$, the equation system (13) will lead to $\tau_{t+1}^{\frac{t}{2}} = 0$ or $\tau_{t+1}'^{\frac{t}{2}} = 0$.*

Proof. We may as well let $r_1 > r_2$. With the input difference $\Delta_1^{forward} = (0, \cdots, 0, \alpha_1)$, it will activate the S-box sequentially as below

$$\alpha_1 \xrightarrow{S_2^{\frac{t}{2}}} \tau_2^{\frac{t}{2}} \xrightarrow{S_3^{\frac{t}{2}}} \cdots \xrightarrow{S_{t+1}^{\frac{t}{2}}} \tau_{t+1}^{\frac{t}{2}}.$$

Then after t rounds, it has $\Delta_{1+t}^{forward} = (*, \cdots, *, \alpha_1)$. For the following two rounds, it has

$$\Delta_{1+t+1}^{forward} = (*, \cdots, *, \alpha_1 + \tau_{t+1}^{\frac{t}{2}}, *),$$

$$\Delta_{1+t+2}^{forward} = (*, \cdots, *, \alpha_1 + \tau_{t+1}^{\frac{t}{2}}, *, *).$$

As for the output difference $\nabla_1^{backward} = (0, \cdots, 0, \beta_1, 0)$, thus if $r_1+1 \equiv t-r_2+2 \bmod t$, that is $r_1 + r_2 \equiv 1 \bmod t$, the block difference $\alpha_1 + \tau_{t+1}^{\frac{t}{2}}$ and β_1 will coincide. Considering $\alpha_1 = \beta_1$, it must have $\tau_{t+1}^{\frac{t}{2}} = 0$. Similarly, if $r_1 < r_2$, it will lead to $\tau_{t+1}^{'\frac{t}{2}} = 0$. □

4.3.3 Transformation from IDC to ZC and INT of GMiMC$_{\texttt{Nyb}}$

We first reveal an equivalent relation between dual structure and inverse structure of $\mathcal{E}_{Nyb}$, which is given in Lemma 12.

Lemma 12. *The inverse structure $\mathcal{E}_{Nyb}^{-1}$ and dual structure $\mathcal{E}_{Nyb}^{\perp}$ are equivalent.*

Proof. For the Nyber's generalized Feistel structure $\mathcal{E}_{Nyb}$, following the similar representation in [BMT13], its i-th round function can be represented by using $t \times t$ matrix $P \circ F_i$ as below

$$F_i = \begin{pmatrix} \begin{smallmatrix} 1 & 0 \\ S_i^1 & 1 \end{smallmatrix} & & & & \\ & \begin{smallmatrix} 1 & 0 \\ S_i^2 & 1 \end{smallmatrix} & & & \\ & & \ddots & & \\ & & & \begin{smallmatrix} 1 & 0 \\ S_i^{\frac{t}{2}} & 1 \end{smallmatrix} \end{pmatrix}, P = \begin{pmatrix} & 1 & & \\ & & \ddots & \\ & & & 1 \\ 1 & & & \end{pmatrix}.$$

where $F_i^{-1} = F_i$ and $P^T = P^{-1}$. For the dual structure of $\mathcal{E}_{Nyb}$, we have

$$F_i^T = \begin{pmatrix} \begin{smallmatrix} 1 & S_i^1 \\ 0 & 1 \end{smallmatrix} & & & & \\ & \begin{smallmatrix} 1 & S_i^2 \\ 0 & 1 \end{smallmatrix} & & & \\ & & \ddots & & \\ & & & \begin{smallmatrix} 1 & S_i^{\frac{t}{2}} \\ 0 & 1 \end{smallmatrix} \end{pmatrix}, (P^{-1})^T = P = \begin{pmatrix} & 1 & & \\ & & \ddots & \\ & & & 1 \\ 1 & & & \end{pmatrix}.$$

Now, consider two permutation π_1 and π_2 as below

$$\pi_1 = \begin{pmatrix} \begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix} & & & & \\ & \begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix} & & & \\ & & \ddots & & \\ & & & \begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix} \end{pmatrix}, \pi_2 = \begin{pmatrix} & & & \begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix} \\ & & \begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix} & \\ & \iddots & & \\ \begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix} & & & \end{pmatrix},$$

where $\pi_1 = \pi_1^{-1} = \pi_1^T$ and $\pi_2 = \pi_2^{-1} = \pi_2^T$. As it has that $F_i^T = \pi_1^{-1} F_i \pi_1$ and $(P^{-1})^T = \pi_2^{-1} P \pi_2$, we let

$$\Pi = \pi_1 \pi_2 = \begin{pmatrix} & & & 1 \\ & & \iddots & \\ & 1 & & \\ 1 & & & \end{pmatrix},$$

where $\Pi = \Pi^{-1}$ and $\Pi P \Pi = P^T = P^{-1}$.

Now, for n rounds $\mathcal{E}_{Nyb}^{\perp}$, it can be expanded as

$$\begin{aligned}(P^{-1})^T \circ F_{n-1}^T \cdots (P^{-1})^T \circ F_0^T &= \pi_2^{-1} P \pi_2 \circ \pi_1^{-1} F_{n-1} \pi_1 \cdots \pi_2^{-1} P \pi_2 \circ \pi_1^{-1} F_0 \pi_1 \\ &= \pi_2^{-1} P \circ \Pi \circ F_{n-1} \circ \Pi \cdots \circ \Pi \circ P \circ \Pi \circ F_0 \pi_1 \\ &= \pi_2^{-1} P \circ \Pi \circ F_{n-1} \circ P^{-1} \cdots F_1 \circ P^{-1} \circ F_0 \pi_1 \\ &= \pi_2^{-1} P \circ \Pi \circ (F_{n-1}^{-1} \circ P^{-1} \cdots F_0^{-1} \circ P^{-1}) \circ P \pi_1 \\ &= A \circ (F_{n-1}^{-1} \circ P^{-1} \cdots F_0^{-1} \circ P^{-1}) \circ B.\end{aligned}$$

where $A = \pi_2^{-1} P \Pi$ and $B = P \pi_1$. Thus, it has that $\mathcal{E}_{Nyb}^{-1}$ is equivalent to $\mathcal{E}_{Nyb}^{\perp}$. □

With this underlying equivalent relation of these structures for $\mathcal{E}_{Nyb}$, according to Theorem 5, we know there is a one-to-one correspondence between IDC and ZC for $\mathcal{E}_{Nyb}$. Then, we can transform two IDCs presented above to its corresponding ZCs of GMiMC$_{\texttt{Nyb}}$.

For $(2t-1)$-round IDC, its input difference is $\alpha = (0, \cdots, 0, \alpha_1)$ and the output difference is $\beta = (\beta_1, 0, \cdots, 0, 0)$, where α_1, β_1 are non-zero. So, $\alpha \to \beta$ is a ZC for the dual structure $\mathcal{E}_{Nyb}^{\perp}$. As proved in Lemma 12, the dual structure $\mathcal{E}_{Nyb}^{\perp}$ is equivalent to the inverse structure $\mathcal{E}_{Nyb}^{-1}$, that is $\mathcal{E}_{Nyb}^{\perp} = A \circ \mathcal{E}_{Nyb}^{-1} \circ B$, where the matrix representation for A and B are given as below

$$A = B = \begin{pmatrix} & & & 1 \\ & & \iddots & \\ & 1 & & \\ 1 & & & \end{pmatrix}.$$

Thus, $\alpha \to \beta$ is a ZC for the equivalent inverse structure $A \circ \mathcal{E}_{Nyb}^{-1} \circ B$. After passing the linear transformation A^{-1} and B, $\alpha' \to \beta'$ is a ZC for the inverse structure $\mathcal{E}_{Nyb}^{-1}$

$$\alpha' = B \circ \alpha = (0, \alpha_1, 0, \cdots, 0) \text{ and } \beta' = A^{-1} \circ \beta = (\beta_1, 0, \cdots, 0).$$

Naturally, $\beta' \to \alpha'$ is a ZC for GMiMC$_{\texttt{Nyb}}$. That is, $(b_1, 0, \cdots, 0) \to (0, a_1, 0, \cdots, 0)$ is a $(2t-1)$-round ZC for GMiMC$_{\texttt{Nyb}}$, where $a_1, b_1 \neq 0$. Similarly, based on $(2t+1)$-round IDC for GMiMC$_{\texttt{Nyb}}$, we know that $(0, 0, b_1, 0, \cdots, 0) \to (0, a_1, 0, \cdots, 0)$ is a $(2t+1)$-round ZC for GMiMC$_{\texttt{Nyb}}$, where $a_1, b_1 \neq 0$ and $a_1 = b_1$.

As for integral, we can obtain the transformed $(2t-1)$-round INT for GMiMC$_{\texttt{Nyb}}$ by using Theorem 10, let $V = \{(x, 0, \cdots, 0) | x \in \mathbb{F}_p\}$, if the input space is $V^{\perp}$, then the output is balanced on $(0, a_1, 0, \cdots, 0) \cdot$GMiMC$_{\texttt{Nyb}}$.

4.4 Equation-based Method for GMiMC$_{\texttt{mrf}}$

For another balanced Feistel construction GMiMC$_{\texttt{mrf}}$ with t branches, its full diffusion rounds is $\Lambda(t) = 2\lceil \log_2(t) \rceil$, which is achieved by its Multi-Rotating round function. In this section, we only focus on the case where t is exactly power of two, and the round index of the distinguisher starts from 1 (due to the different rotation constant s_r for each round). Similarly, its IDC consists of the forward $r_1 (\Lambda(t) - 2 \leq r_1 \leq \Lambda(t) + 2)$ rounds and backward $r_2 (\Lambda(t) - 2 \leq r_2 \leq \Lambda(t) + 2)$ rounds.

For the round function of GMiMC$_{\texttt{mrf}}$, the input difference $\Delta_1^{forward} = (0, \cdots, 0, \alpha_1)$ can pass r_1 rounds forwards to the output difference $\Delta_{1+r_1}^{forward}$ as below

$$(*, \cdots, *, \underbrace{\alpha_1 + \sum_{r+=2}^{1 \leq r \leq r_1} \tau_r^{\frac{t}{2}}}_{(r_1 \bmod 2)\frac{t}{2}+1}, *, \cdots, *).$$

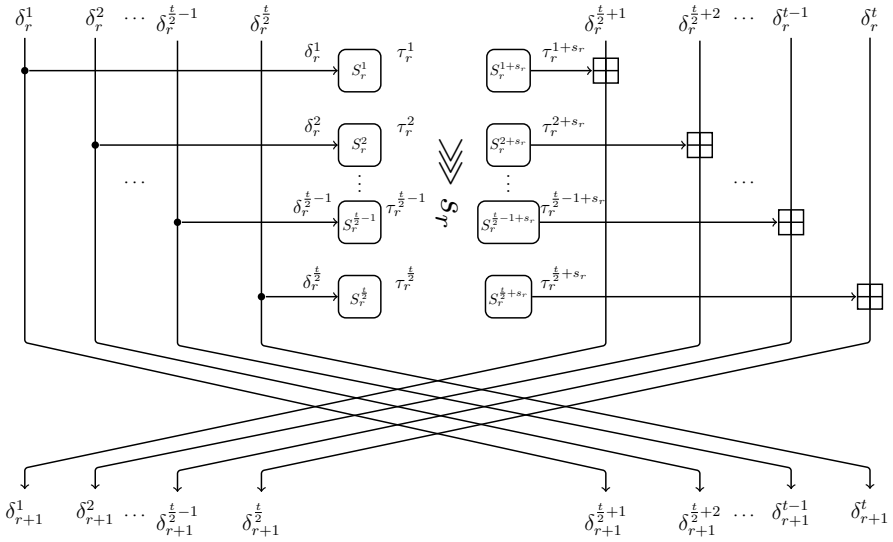


Figure 15: Differential of GMiMC$_{\texttt{mrf}}$'s r-th round function with t branches.

Similarly, the output difference $\nabla_1^{backward} = (0, \cdots, 0, \beta_1, \underbrace{0, \cdots, 0}_{\frac{t}{2}})$ passes r_2 rounds backwards to $\nabla_{r_2+1}^{backward}$ as below

$$(*, \cdots, *, \underbrace{\beta_1 - \sum_{r+=2}^{1 \le r \le r_2} {\tau_r}'^{\frac{t}{2}}, *, \cdots, *}_{((r_2+1) \bmod 2)\frac{t}{2}+1}).$$

Then for a valid differential trail, the output difference and the input difference meeting in the middle should be equal, and it has the following

$$\Delta_{1+r_1}^{forward} = \nabla_{r_2+1}^{backward}. \tag{14}$$

Naturally, if we find some contradictions in the above equation systems, it will lead to an IDC with $(r_1 + r_2)$ rounds for GMiMC$_{\texttt{mrf}}$. Before presenting the IDC of GMiMC$_{\texttt{mrf}}$, some properties are first prepared for constructing its IDC.

Property 1. For the rotations of GMiMC$_{\texttt{mrf}}$, we have the sum of rotations of consecutive $\Lambda(t) - 2$ or $\Lambda(t) - 1$ rounds (starting from round 1)

$$RotSum(\Lambda(t) - 2) = \sum_{1 \le i \le \Lambda(t)-2} s_i = 2^{\frac{\Lambda(t)}{2}-1} - 1 = \frac{t}{2} - 1,$$

$$RotSum(\Lambda(t) - 1) = \sum_{1 \le i \le \Lambda(t)-1} s_i = 2^{\frac{\Lambda(t)}{2}-1} - 1 = \frac{t}{2} - 1.$$

Similarly, the sum of rotations of consecutive $\Lambda(t)$ or $\Lambda(t) + 1$ rounds (starting from round 1)

$$RotSum(\Lambda(t)) = RotSum(\Lambda(t) + 1) = 2^{\frac{\Lambda(t)}{2}-1} = \frac{t}{2}.$$

Property 2. For GMiMC$_{\texttt{mrf}}$ with input difference $\Delta_1^{forward} = (0, \cdots, 0, \alpha_1)$, within $\Lambda(t)$ rounds forwards, the block with difference α_1 will not be involved with any other non-zero differences.

Property 3. For GMiMC$_{\texttt{mrf}}$ with output difference $\nabla_1^{backward} = (0, \cdots, \beta_1, \underbrace{0, \cdots, 0}_{\frac{t}{2}})$, within $\Lambda(t)$ rounds backwards, the block with difference β_1 will not be involved with any other non-zero differences.

Property 4. For GMiMC$_{\texttt{mrf}}$ with input difference $\Delta_1^{forward} = (0, \cdots, 0, \alpha_1)$, after $\Lambda(t)-2$ rounds forwards, the $\frac{t}{2}$-th block from left is with zero difference.

Property 5. For GMiMC$_{\texttt{mrf}}$ with output difference $\nabla_1^{backward} = (0, \cdots, \beta_1, \underbrace{0, \cdots, 0}_{\frac{t}{2}})$, after $\Lambda(t) - 2$ rounds backwards, the rightmost block is with zero difference (starting from rotation constant with 0)

Property 6. For GMiMC$_{\texttt{mrf}}$ with input difference $\Delta_1^{forward} = (0, \cdots, 0, \alpha_1)$, after $\Lambda(t)-1$ rounds forwards, the rightmost block is with zero difference.

Property 7. For GMiMC$_{\texttt{mrf}}$ with output difference $\nabla_1^{backward} = (0, \cdots, \beta_1, \underbrace{0, \cdots, 0}_{\frac{t}{2}})$, after $\Lambda(t) - 1$ rounds backwards, the $\frac{t}{2}$-th block from left is with zero difference (starting from rotation constant with 0).

Property 8. For GMiMC$_{\texttt{mrf}}$ with input difference $\Delta_1^{forward} = (0, \cdots, 0, \alpha_1)$, the following $\Lambda(t) + 1$ rounds, it will activate the S-box as below

$$\alpha_1 \xrightarrow{S_2^{\frac{t}{2}+s_1}} \tau_2^{\frac{t}{2}+s_1} \cdots \xrightarrow{S_{\Lambda(t)}^{\frac{t}{2}+RotSum(\Lambda(t)-1)}} \tau_{\Lambda(t)}^{\frac{t}{2}-1} \xrightarrow{S_{\Lambda(t)+1}^{\frac{t}{2}+RotSum(\Lambda(t))}} \tau_{\Lambda(t)+1}^{\frac{t}{2}}.$$

Property 9. For GMiMC$_{\texttt{mrf}}$ with output difference $\nabla_1^{backward} = (0, \cdots, \beta_1, \underbrace{0, \cdots, 0}_{\frac{t}{2}})$, the following $\Lambda(t) + 1$ rounds, it will activate the S-box as below

$$\beta_1 \xrightarrow{S_2^{\frac{t}{2}+s'_1}} \tau'^{\frac{t}{2}+s'_1}_2 \cdots \rightarrow \tau'^{\frac{t}{2}+RotSum(\Lambda(t))-s'_{\Lambda(t)}}_{\Lambda(t)} \xrightarrow{S_{\Lambda(t)+1}^{\frac{t}{2}+RotSum(\Lambda(t))}} \tau'^{\frac{t}{2}}_{\Lambda(t)+1}.$$

From Property 2 to Property 9, all these properties can be deduced from the special diffusion properties of Multi-Rotating round function of GMiMC$_{\texttt{mrf}}$, its fast diffusion is contributed by a sequence of the selected rotations for each round, which leads to full diffusion after $\Lambda(t)$ rounds. For more details of this Multi-Rotating round function, we refer the reader to the design paper [AGP+19b, Section 2.1.4]. Now based on these properties, IDC of GMiMC$_{\texttt{mrf}}$ will be constructed as follws.

4.4.1 IDC of GMiMC$_{\texttt{mrf}}$ with $2\Lambda(t) + 1$ Rounds

Lemma 13. *For the power-of-two branch t, if $r_1 + r_2 = 2\Lambda(t) + 1$ and $\alpha_1 = \beta_1$ where $\Lambda(t)-2 \leq r_1, r_2 \leq \Lambda(t)+2$, the equation system (14) will lead to $\tau_{\Lambda(t)+1}^{\frac{t}{2}} = 0$ or $\tau'^{\frac{t}{2}}_{\Lambda(t)+1} = 0$.*

Proof. We may as well let $r_1 > r_2$. Due to the condition on the number of rounds $\Lambda(t) - 2 \leq r_1, r_2 \leq \Lambda(t) + 2$, then $r_1 = \Lambda(t) + 1$ or $\Lambda(t) + 2$. According to Property 2, after $\Lambda(t)$ rounds, we have $\Delta_{1+\Lambda(t)}^{forward} = (*, \cdots, *, \alpha_1)$. Then according to Property 8, for the following two rounds, it has

$$\Delta_{1+\Lambda(t)+1}^{forward} = (*, \cdots, *, \alpha_1 + \tau_{\Lambda(t)+1}^{\frac{t}{2}}),$$

$$\Delta_{1+\Lambda(t)+2}^{forward} = (*, \cdots, *, \alpha_1 + \tau_{\Lambda(t)+1}^{\frac{t}{2}}, \underbrace{*, \cdots, *}_{\frac{t}{2}}).$$

For the output difference $\nabla_1^{backward} = (0, \cdots, 0, \beta_1, \underbrace{0, \cdots, 0}_{\frac{t}{2}})$, according to Property 3, it has $\nabla_{r_2+1}^{backward}$

$$(*, \cdots, *, \quad \underbrace{\beta_1, *, \cdots, *}_{((r_2+1) \bmod 2)\frac{t}{2}+1} \quad).$$

As $r_1 + r_2 = 2\Lambda(t) + 1$ is odd, the block differences $\alpha_1 + \tau_{\Lambda(t)+1}^{\frac{t}{2}}$ and β_1 will coincide. Considering $\alpha_1 = \beta_1$, it will lead to $\tau_{\Lambda(t)+1}^{\frac{t}{2}} = 0$. Similarly, if $r_1 < r_2$, according to Property 9, it will lead to $\tau_{\Lambda(t)+1}^{'\frac{t}{2}} = 0$. □

For any nontrivial differential trail with input difference $(0, \cdots, 0, \alpha_1)$ and output difference $(0, \cdots, 0, \beta_1, \underbrace{0, \cdots, 0}_{\frac{t}{2}})$, where $\alpha_1 = \beta_1 \neq 0$. An IDC with $(2\Lambda(t) + 1)$ rounds of GMiMC$_{\texttt{mrf}}$ is obtained.

4.4.2 IDC of GMiMC$_{\texttt{mrf}}$ with $2\Lambda(t) - 1$ Rounds

Lemma 14. *For the power-of-two branch t, if $r_1 + r_2 = 2\Lambda(t) - 1$ where $\Lambda(t) - 2 \leq r_1, r_2 \leq \Lambda(t) + 1$, the equation system (14) will lead to $\tau_{\Lambda(t)}^{\frac{t}{2}-1} = 0$ or $\tau_{\Lambda(t)}^{'\frac{t}{2}-1} = 0$.*

Proof. We may as well let $r_1 > r_2$. Due to the condition on the number of rounds $\Lambda(t) - 2 \leq r_1, r_2 \leq \Lambda(t) + 1$, then $r_1 = \Lambda(t)$ or $\Lambda(t) + 1$. With input difference $\Delta_1^{forward} = (0, \cdots, 0, \alpha_1)$, it will activate the S-box sequentially as below

$$\alpha_1 \xrightarrow{S_2^{\frac{t}{2}}} \tau_2^{\frac{t}{2}} \xrightarrow{S_3^{\frac{t}{2}+s_2}} \tau_3^{\frac{t}{2}+s_2} \xrightarrow{S_4^{\frac{t}{2}+s_2+s_3}} \cdots \xrightarrow{S_{\Lambda(t)}^{\frac{t}{2}+s_2+s_3+\cdots s_{\Lambda(t)-1}}} \tau_{\Lambda(t)}^{\frac{t}{2}+s_2+s_3+\cdots s_{\Lambda(t)-1}},$$

still according to Property 1, it has $\tau_{\Lambda(t)}^{\frac{t}{2}+s_2+\cdots+s_{\Lambda(t)-1}} = \tau_{\Lambda(t)}^{\frac{t}{2}-1}$. Similarly, we have $\Delta_{1+r_1}^{forward}$

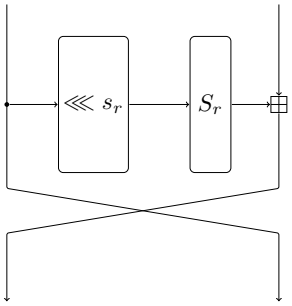
$$(*, \cdots, \tau_{\Lambda(t)}^{\frac{t}{2}-1}, \underbrace{*, \cdots, \alpha_1}_{\frac{t}{2}}) \text{ or } (*, \cdots, *, \underbrace{*, \cdots, \tau_{\Lambda(t)}^{\frac{t}{2}-1}}_{\frac{t}{2}}).$$

Then according to Property 5 and 7, for the output difference $\nabla_1^{backward}$ with $r_1 + r_2 = 2\Lambda(t) - 1$, it has $\nabla_{r_2+1}^{backward}$

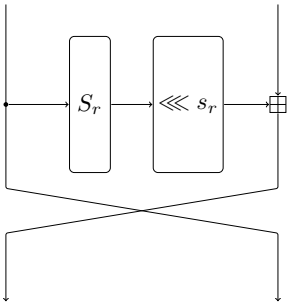
$$(*, \cdots, \beta_1, \underbrace{*, \cdots, 0}_{\frac{t}{2}}) \text{ or } (*, \cdots, 0, \underbrace{*, \cdots, \beta_1}_{\frac{t}{2}}).$$

Then the block differences $\tau_{\Lambda(t)}^{\frac{t}{2}-1}$ and 0 will coincide, which means $\tau_{\Lambda(t)}^{\frac{t}{2}-1} = 0$. Similarly, if $r_1 < r_2$, according to Property 4 and 6, it will lead to $\tau_{\Lambda(t)}^{'\frac{t}{2}-1} = 0$. □

Naturally, for any nontrivial differential trail with input difference $(0, 0, \cdots, 0, \alpha_1)$ and output difference $(0, \cdots, 0, \beta_1, \underbrace{0, \cdots, 0}_{\frac{t}{2}})$, where $\alpha_1, \beta_1 \neq 0$, from Lemma 14, an IDC with $(2\Lambda(t) - 1)$ rounds of GMiMC$_{\texttt{mrf}}$ is obtained.



(a) The round function of $\mathcal{E}^{\perp}_{mrf}$.



(b) Equivalent structure of $\mathcal{E}^{\perp}_{mrf}$.

Figure 16: Structure of $\mathcal{E}^{\perp}_{mrf}$ and its equivalent.

4.4.3 IDC of the Dual Structure $\mathcal{E}^{\perp}_{mrf}$

Now, we consider the IDC of the dual structure $\mathcal{E}^{\perp}_{mrf}$, which is depicted in Figure 16(a), and its equivalent structure is depicted in Figure 16(b). As can be observed in $\mathcal{E}^{\perp}_{mrf}$, the rotation in each round is reversed, that is $-s_r$.

Lemma 15. *For $\mathcal{E}^{\perp}_{mrf}$ with the power-of-two branch t, if $r_1 + r_2 = 2\Lambda(t) + 1$ and $\alpha_1 = \beta_1$ where $\Lambda(t) - 2 \leq r_1, r_2 \leq \Lambda(t) + 2$, the equation system (14) will lead to $\tau^{\frac{t}{2}}_{\Lambda(t)+1} = 0$ or $\tau'^{\frac{t}{2}}_{\Lambda(t)+1} = 0$.*

Proof. As Property 2, 3, 8 and 9 still hold when having the reversed rotation $-s_r$, the same result can be deduced like in Lemma 13. □

Lemma 16. *For $\mathcal{E}^{\perp}_{mrf}$ with the power-of-two branch t, if $r_1 + r_2 = 2\Lambda(t) - 1$ where $\Lambda(t) - 2 \leq r_1, r_2 \leq \Lambda(t) + 1$, the equation system (14) will lead to $\tau^{1}_{\Lambda(t)} = 0$ or $\tau'^{1}_{\Lambda(t)} = 0$.*

Proof. Property 4, 5, 6 and 7 still hold when having the reversed rotation $-s_r$. While Property 1 will be slightly different, for inverse rotation $-s_r$, it has $RotSum(\Lambda(t) - 1) = RotSum(\Lambda(t) - 2) = 1 - \frac{t}{2}$ and just changes the index. So, the similar result can be deduced like in Lemma 14. □

4.4.4 Transformation from IDC of $\mathcal{E}^{\perp}_{mrf}$ to ZC and INT of $\mathcal{E}_{mrf}$

For two IDCs of $\mathcal{E}^{\perp}_{mrf}$ presented above, by the transformations depicted in Figure 17, we obtain two ZCs of GMiMC$_{\texttt{mrf}}$ as below

- $(2\Lambda(t)-1)$ rounds: with input mask $(0, \cdots, 0, a_1, \underbrace{0, \cdots, 0}_{\frac{t}{2}})$ and output mask $(0, \cdots, 0, b_1)$, where $a_1, b_1 \neq 0$.
- $(2\Lambda(t)+1)$ rounds: with input mask $(0, \cdots, 0, a_1, \underbrace{0, \cdots, 0}_{\frac{t}{2}})$ and output mask $(0, \cdots, 0, b_1)$, where $a_1 = b_1 \neq 0$.

Then based on Theorem 9, an INT of $\mathcal{E}_{mrf}$ with $(2\Lambda(t) - 1)$ rounds is obtained. The the input space is $V^{\perp}$ and the output is balanced on $(0, \cdots, 0, b_1)\cdot$GMiMC$_{\texttt{mrf}}$, where $V = \{(0, \cdots, 0, x, \underbrace{0, \cdots, 0}_{\frac{t}{2}}) | x \in \mathbb{F}_p\}$.

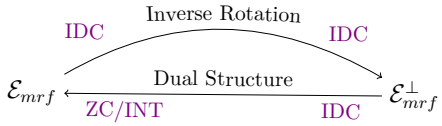


Figure 17: Transformation from IDC of $\mathcal{E}^{\perp}_{mrf}$ to ZC/INT of $\mathcal{E}_{mrf}$.

Experiments: All these constructed or transformed INTs and ZCs for all GMiMC constructions presented above are verified by experiments on small instances. The details of the experiments are given in Appendix C and our codes are provided at https://github.com/csy1234/Links_IDC_ZC_INT.

5 Conclusion

In this paper, we have established the comprehesive links between impossible differential, zero-correlation linear and integral cryptanalysis over the prime field $\mathbb{F}_p$, for the very first time. The links between zero-correlation linear and integral cryptanalysis are also proved in an alternative way, through which we find that the independent conditions of the input and output masks (differences) cannot be removed when deriving an integral distinguisher from a zero-correlation linear hull (impossible differential) over $\mathbb{F}_p$, this exhibits a difference of these cryptanalytic methods between $\mathbb{F}_p$ and $\mathbb{F}_2^n$.

To showcase our refined links, we apply to GMiMC and obtain different type of improved distinguishers for all GMiMC constructions, from which the gaps of symmetric cryptanalytic methods between $\mathbb{F}_p$ and $\mathbb{F}_2^n$ are also demonstrated in terms of attacked rounds, even distinguishers with an *arbitrary number of rounds* for some special and limited cases. The establishment of the theories over $\mathbb{F}_p$ behind these links, and properties identified (be it similar or different) will bring clearer and easier understanding of security of MPC/FHE/ZK-friendly symmetric-key primitives, which could facilitate the future design and cryptanalysis.

Further discussions. Considering only the characteristic p is relevant and the isomorphism from $\mathbb{F}_{p^t}$ to $\mathbb{F}_p^t$, for the proposed works in this paper, there is possible generalization to $\mathbb{F}_q$ where $q = p^t$, which could be used for the potential MPC/FHE/ZK-friendly designs over $\mathbb{F}_q$ in the future. Secondly, the statistical cryptanalytic method is still missing for zero-correlation linear cryptanalysis over $\mathbb{F}_p$, thus more dedicated statistical model should be developed to evaluate the detailed complexity of the attack. Thirdly, according to our proposed links over $\mathbb{F}_p$, an integral distinguisher arising from low-degree S-box (i.e. zero-sum property) does not imply any impossible differential or zero-correlation linear hull, this also has been observed in [SLR+15] for the links over $\mathbb{F}_2^n$, which still needs to be investigated further.

Acknowledgements: The authors would like to thank all the anonymous reviewers for their helpful and detailed comments. We also want to thank Christian Rechberger for his important comments on this work.

This research is supported by the National Research Foundation, Singapore under its Strategic Capability Research Centres Funding Initiative, the Nanyang Technological University in Singapore under Start-up Grant 04INS000397C230, and Ministry of Education in Singapore under Grants RG91/20, the National Key Research and Development Program of China (Grant No. 2018YFA0704702), the National Natural Science Foundation of China (Grant No. 62032014), the Major Basic Research Project of Natural Science Foundation of Shandong Province, China (Grant No. ZR202010220025), the National Key R&D Program

of China (Grant No. 2022YFB2701700), Shandong Provincial Natural Science Foundation (Grant No. ZR2020MF053) and the National Natural Science Foundation of China (Grant No. 62002202). Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not reflect the views of National Research Foundation, Singapore.

References

[AAB+20] Abdelrahaman Aly, Tomer Ashur, Eli Ben-Sasson, Siemen Dhooghe, and Alan Szepieniec. Design of symmetric-key primitives for advanced cryptographic protocols. *IACR Trans. Symmetric Cryptol.*, 2020(3):1–45, 2020.

[AD18] Tomer Ashur and Siemen Dhooghe. MARVELlous: a STARK-friendly family of cryptographic primitives. *IACR Cryptol. ePrint Arch.*, 2018:1098, 2018.

[AGP+19a] Martin R. Albrecht, Lorenzo Grassi, Léo Perrin, Sebastian Ramacher, Christian Rechberger, Dragos Rotaru, Arnab Roy, and Markus Schofnegger. Feistel Structures for MPC, and More. In Kazue Sako, Steve A. Schneider, and Peter Y. A. Ryan, editors, *ESORICS 2019, Part II*, volume 11736 of *LNCS*, pages 151–171. Springer, 2019.

[AGP+19b] Martin R. Albrecht, Lorenzo Grassi, Leo Perrin, Sebastian Ramacher, Christian Rechberger, Dragos Rotaru, Arnab Roy, and Markus Schofnegger. Feistel structures for mpc, and more. Cryptology ePrint Archive, Paper 2019/397, 2019.

[AGR+16] Martin R. Albrecht, Lorenzo Grassi, Christian Rechberger, Arnab Roy, and Tyge Tiessen. MiMC: Efficient encryption and cryptographic hashing with minimal multiplicative complexity. In Jung Hee Cheon and Tsuyoshi Takagi, editors, *ASIACRYPT 2016, Part I*, volume 10031 of *LNCS*, pages 191–219, 2016.

[AP11] Andrea Agnesse and Marco Pedicini. Cube attack in finite fields of higher order. In *Ninth Australasian Information Security Conference, AISC 2011, Perth, Australia, January 2011*, pages 9–14, 2011.

[ARS+15] Martin R. Albrecht, Christian Rechberger, Thomas Schneider, Tyge Tiessen, and Michael Zohner. Ciphers for MPC and FHE. In Elisabeth Oswald and Marc Fischlin, editors, *EUROCRYPT 2015, Part I*, volume 9056 of *LNCS*, pages 430–454. Springer, 2015.

[BBC+22] Clémence Bouvier, Pierre Briaud, Pyrros Chaidos, Léo Perrin, Robin Salen, Vesselin Velichkov, and Danny Willems. New design techniques for efficient arithmetization-oriented hash functions: Anemoi permutations and jive compression mode. Cryptology ePrint Archive, Paper 2022/840, 2022.

[BBW14] Céline Blondeau, Andrey Bogdanov, and Meiqin Wang. On the (in)equivalence of impossible differential and zero-correlation distinguishers for feistel- and skipjack-type ciphers. In *ACNS 2014*, pages 271–288, 2014.

[BCD+20] Tim Beyne, Anne Canteaut, Itai Dinur, Maria Eichlseder, Gregor Leander, Gaëtan Leurent, María Naya-Plasencia, Léo Perrin, Yu Sasaki, Yosuke Todo, and Friedrich Wiemer. Out of oddity - new cryptanalytic techniques against symmetric primitives optimized for integrity proof systems. In Daniele Micciancio and Thomas Ristenpart, editors, *Advances in Cryptology - CRYPTO 2020, Part III*, volume 12172 of *LNCS*, pages 299–328. Springer, 2020.

[BEF19] Dan Boneh, Saba Eskandarian, and Ben Fisch. Post-quantum EPID signatures from symmetric primitives. In *Topics in Cryptology - CT-RSA 2019*, pages 251–271, 2019.

[Bey21] Tim Beyne. A geometric approach to linear cryptanalysis. In *Advances in Cryptology - ASIACRYPT 2021, Part I*, pages 36–66, 2021.

[BL22] Tim Beyne and Yunwen Liu. Truncated differential attacks on contracting feistel ciphers. *IACR Trans. Symmetric Cryptol.*, 2022(2):141–160, 2022.

[BLNW12] Andrey Bogdanov, Gregor Leander, Kaisa Nyberg, and Meiqin Wang. Integral and multidimensional linear distinguishers with correlation zero. In *ASIACRYPT 2012*, pages 244–261, 2012.

[BMT13] Thierry P. Berger, Marine Minier, and Gaël Thomas. Extended generalized feistel networks using matrix representation. In *SAC 2013*, pages 289–305, 2013.

[BN13] Céline Blondeau and Kaisa Nyberg. New links between differential and linear cryptanalysis. In Thomas Johansson and Phong Q. Nguyen, editors, *Advances in Cryptology - EUROCRYPT 2013, 32nd Annual International Conference on the Theory and Applications of Cryptographic Techniques, Athens, Greece, May 26-30, 2013. Proceedings*, volume 7881 of *Lecture Notes in Computer Science*, pages 388–404. Springer, 2013.

[BN14] Céline Blondeau and Kaisa Nyberg. Links between truncated differential and multidimensional linear properties of block ciphers and underlying attack complexities. In *EUROCRYPT 2014*, pages 165–182, 2014.

[Bon19] Xavier Bonnetain. Collisions on feistel-mimc and univariate gmimc. Cryptology ePrint Archive, Paper 2019/951, 2019.

[BR14] Andrey Bogdanov and Vincent Rijmen. Linear hulls with correlation zero and linear cryptanalysis of block ciphers. *Des. Codes Cryptogr.*, 70(3):369–383, 2014.

[BSV07] Thomas Baignères, Jacques Stern, and Serge Vaudenay. Linear cryptanalysis of non binary ciphers. In Carlisle M. Adams, Ali Miri, and Michael J. Wiener, editors, *SAC 2007*, volume 4876 of *LNCS*, pages 184–211. Springer, 2007.

[CCF+18] Anne Canteaut, Sergiu Carpov, Caroline Fontaine, Tancrède Lepoint, María Naya-Plasencia, Pascal Paillier, and Renaud Sirdey. Stream ciphers: A practical solution for efficient homomorphic-ciphertext compression. *J. Cryptol.*, 31(3):885–916, 2018.

[CDG+17] Melissa Chase, David Derler, Steven Goldfeder, Claudio Orlandi, Sebastian Ramacher, Christian Rechberger, Daniel Slamanig, and Greg Zaverucha. Post-quantum zero-knowledge and signatures from symmetric-key primitives. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security, CCS 2017*, pages 1825–1842, 2017.

[CIR22] Carlos Cid, John Petter Indrøy, and Håvard Raddum. FASTA - A stream cipher for fast FHE evaluation. In *Topics in Cryptology - CT-RSA 2022*, pages 451–483, 2022.

[CV94] Florent Chabaud and Serge Vaudenay. Links between differential and linear cryptanalysis. In Alfredo De Santis, editor, *Advances in Cryptology - EUROCRYPT '94, Workshop on the Theory and Application of Cryptographic Techniques, Perugia, Italy, May 9-12, 1994, Proceedings*, volume 950 of *Lecture Notes in Computer Science*, pages 356–365. Springer, 1994.

[DEG+18] Christoph Dobraunig, Maria Eichlseder, Lorenzo Grassi, Virginie Lallemand, Gregor Leander, Eik List, Florian Mendel, and Christian Rechberger. Rasta: A cipher with low anddepth and few ands per bit. In *Advances in Cryptology - CRYPTO 2018, Part I*, pages 662–692, 2018.

[DGGK21] Christoph Dobraunig, Lorenzo Grassi, Anna Guinet, and Daniël Kuijsters. Ciminion: Symmetric encryption based on toffoli-gates over large finite fields. In Anne Canteaut and François-Xavier Standaert, editors, *Advances in Cryptology - EUROCRYPT 2021, Part II*, volume 12697 of *LNCS*, pages 3–34. Springer, 2021.

[DGH+21] Christoph Dobraunig, Lorenzo Grassi, Lukas Helminger, Christian Rechberger, Markus Schofnegger, and Roman Walch. Pasta: A case for hybrid homomorphic encryption. *IACR Cryptol. ePrint Arch.*, page 731, 2021.

[GHR+22] Lorenzo Grassi, Yonglin Hao, Christian Rechberger, Markus Schofnegger, Roman Walch, and Qingju Wang. Horst meets fluid-spn: Griffin for zero-knowledge applications. Cryptology ePrint Archive, Paper 2022/403, 2022.

[GKL+22] Lorenzo Grassi, Dmitry Khovratovich, Reinhard Lüftenegger, Christian Rechberger, Markus Schofnegger, and Roman Walch. Reinforced concrete: A fast hash function for verifiable computation. In *Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security*, CCS 2022, pages 1323–1335, 2022.

[GKR+21] Lorenzo Grassi, Dmitry Khovratovich, Christian Rechberger, Arnab Roy, and Markus Schofnegger. Poseidon: A new hash function for zero-knowledge proof systems. In *30th USENIX Security Symposium, USENIX Security 2021, August 11-13, 2021*, pages 519–535, 2021.

[GLR+20] Lorenzo Grassi, Reinhard Lüftenegger, Christian Rechberger, Dragos Rotaru, and Markus Schofnegger. On a generalization of substitution-permutation networks: The HADES design strategy. In Anne Canteaut and Yuval Ishai, editors, *Advances in Cryptology - EUROCRYPT 2020, Part II*, volume 12106 of *LNCS*, pages 674–704. Springer, 2020.

[KW02] Lars R. Knudsen and David A. Wagner. Integral cryptanalysis. In *Fast Software Encryption, 9th International Workshop, FSE 2002*, pages 112–127, 2002.

[Lai94] Xuejia Lai. Higher order derivatives and differential cryptanalysis. In *Communications and cryptography*, pages 227–233. Springer, 1994.

[Lea11] Gregor Leander. On linear hulls, statistical saturation attacks, PRESENT and a cryptanalysis of PUFFIN. In *EUROCRYPT 2011*, pages 303–322, 2011.

[LLF05] Duo Lei, Chao Li, and Keqin Feng. New observation on camellia. In *Selected Areas in Cryptography, 12th International Workshop, SAC 2005*, pages 51–64, 2005.

[Luc01] Stefan Lucks. The saturation attack - A bait for twofish. In *Fast Software Encryption, 8th International Workshop, FSE 2001 Yokohama, Japan, April 2-4, 2001, Revised Papers*, pages 1–15, 2001.

[MJSC16] Pierrick Méaux, Anthony Journault, François-Xavier Standaert, and Claude Carlet. Towards stream ciphers for efficient FHE with low-noise ciphertexts. In Marc Fischlin and Jean-Sébastien Coron, editors, *EUROCRYPT 2016, Part I*, volume 9665 of *LNCS*, pages 311–343. Springer, 2016.

[SLR+15] Bing Sun, Zhiqiang Liu, Vincent Rijmen, Ruilin Li, Lei Cheng, Qingju Wang, Hoda AlKhzaimi, and Chao Li. Links among impossible differential, integral and zero correlation linear cryptanalysis. In *Advances in Cryptology - CRYPTO 2015, Part I*, pages 95–115, 2015.

A Proof of Theorem 4

Proof. We consider the following two parts.

- We first prove that for a linear hull $(\delta_1^1, \cdots, \delta_1^t) \to (\delta_{r+1}^1, \cdots, \delta_{r+1}^t)$, if $\exists E \in \mathcal{E}_{crf}$ such that $cor((\delta_1^1, \cdots, \delta_1^t) \cdot x - (\delta_{r+1}^1, \cdots, \delta_{r+1}^t) \cdot E(x)) \neq 0$, then $\exists E' \in \mathcal{E}_{erf}$ such that $prob_{E'}((\delta_1^1, \cdots, \delta_1^t), (\delta_{r+1}^1, \cdots, \delta_{r+1}^t)) > 0$.

  As $(\delta_1^1, \cdots, \delta_1^t) \to (\delta_{r+1}^1, \cdots, \delta_{r+1}^t)$ is a linear hull of some $E \in \mathcal{E}_{crf}$ with non-zero correlation, see Figure 4(b), then there must exist a linear characteristic with non-zero correlation, denoted by

  $$(\delta_1^1, \cdots, \delta_1^t) \to \cdots \to (\delta_i^1, \cdots, \delta_i^t) \to \cdots \to (\delta_{r+1}^1, \cdots, \delta_{r+1}^t),$$

  where the input of i-th round can be divided into t branches of $\mathbb{F}_p$ elements, that is $\delta_i = (\delta_i^1, \cdots, \delta_i^t) \in \mathbb{F}_p^t$. Considering this linear characteristic and only one S-box in the round function, the output mask of the non-linear layer is denoted by $\delta_i^1 \in \mathbb{F}_p$ and the input mask of i-th round S-box is denoted by $\beta_i \in \mathbb{F}_p$. Then, based on the propagation rules of linear mask over $\mathbb{F}_p$, it has $\delta_i^j = \beta_i + \delta_{i+1}^{j-1}$ where $1 \leq i \leq r$ and $2 \leq j \leq t$. As for the dual structure of $\mathcal{E}_{crf}$, for any plaintext $(x_1^1, \cdots, x_1^t)$, we construct an r-round cipher $E_r \in \mathcal{E}_{erf}$, such that

  $$E_r(x_1^1, \cdots, x_1^t) - E_r(x_1^1 - \delta_1^1, \cdots, x_1^j - \delta_1^j, \cdots, x_1^t - \delta_1^t) = (\delta_{r+1}^1, \cdots, \delta_{r+1}^t).$$

  When $r = 1$, if $\delta_1^1 = 0$, we can define S'_1 as any possible transformation over $\mathbb{F}_p$, and if $\delta_1^1 \neq 0$, we can define $S'_1(x_1^1) = x_1^1$ and $S'_1(x_1^1 - \delta_1^1) = x_1^1 + \beta_1$. Then for $E_1 \in \mathcal{E}_{erf}$ which adopts such S-box, there will be

  $$\begin{aligned} E_1(x_1^1, \cdots, x_1^t) - E_1(x_1^1 - \delta_1^1, \cdots, x_1^j - \delta_1^j, \cdots, x_1^t - \delta_1^t) &= (\delta_1^2 - \beta_1, \cdots, \delta_1^t - \beta_1, \delta_1^1) \\ &= (\delta_2^1, \cdots, \delta_2^t). \end{aligned}$$

  Suppose that we have constructed E_{r-1} such that

  $$E_{r-1}(x_1^1, \cdots, x_1^t) - E_{r-1}(x_1^1 - \delta_1^1, \cdots, x_1^j - \delta_1^j, \cdots, x_1^t - \delta_1^t) = (\delta_r^1, \cdots, \delta_r^t),$$

  and let $y = (y^1, \cdots, y^t)$ denote the output of $E_{r-1}(x_1^1, \cdots, x_1^t)$. Then in the r-th round, if $\delta_r^1 = 0$, we can define S'_r as any possible transformation over $\mathbb{F}_p$, otherwise, define S'_r as $S'_r(y^1) = y^1$ and $S'_r(y^1 - \delta_r^1) = y^1 + \beta_r$. Therefore, we have

  $$\begin{aligned} E_r(x_1^1, \cdots, x_1^t) - E_r(x_1^1 - \delta_1^1, \cdots, x_1^j - \delta_1^j, \cdots, x_1^t - \delta_1^t) &= (\delta_r^2 - \beta_r, \cdots, \delta_r^t - \beta_r, \delta_r^1) \\ &= (\delta_{r+1}^1, \cdots, \delta_{r+1}^t). \end{aligned}$$

- We secondly prove that for a differential $(\delta_1^1, \cdots, \delta_1^t) \to (\delta_{r+1}^1, \cdots, \delta_{r+1}^t)$, if $\exists E \in \mathcal{E}_{erf}$ such that $prob_E((\delta_1^1, \cdots, \delta_1^t), (\delta_{r+1}^1, \cdots, \delta_{r+1}^t)) > 0$, then $\exists E' \in \mathcal{E}_{crf}$ such that $cor((\delta_1^1, \cdots, \delta_1^t) \cdot x - (\delta_{r+1}^1, \cdots, \delta_{r+1}^t) \cdot E'(x)) \neq 0$.

  As $(\delta_1^1, \cdots, \delta_1^t) \to (\delta_{r+1}^1, \cdots, \delta_{r+1}^t)$ is a differential of some $E \in \mathcal{E}_{erf}$ with non-zero differential probability, also see Figure 4(a), then there must exist a differential characteristic with non-zero probability, denoted as

$$(\delta_1^1, \cdots, \delta_1^t) \to \cdots \to (\delta_i^1, \cdots, \delta_i^t) \cdots \to (\delta_{r+1}^1, \cdots, \delta_{r+1}^t),$$

  where $\delta_i^j \in \mathbb{F}_p (1 \leq i \leq r, 1 \leq j \leq t)$. For this differential characteristic, the input difference of the non-linear layer S_i' is denoted by $\delta_i^1 \in \mathbb{F}_p$. The output difference of S_i' is denoted by $-\beta_i \in \mathbb{F}_p$, then we can connect differences of the expanding part in the round function of $E \in \mathcal{E}_{erf}$, that is $\delta_i^j = \beta_i + \delta_{i+1}^{j-1}$ where $1 \leq i \leq r$ and $2 \leq j \leq t$.

  Considering the following fact: for mask pair (β_i, δ_i^1), where $\delta_i^1 \neq 0$, there always exists an element $a_i \in \mathbb{F}_p$ such that $\beta_i = a_i \delta_i^1$, then for $S_i(x) = a_i x$, we have $cor((\beta_i)^T \cdot x - (\delta_i^1)^T \cdot S_i(x)) = cor((\beta_i - a_i \delta_i^1)^T \cdot x) = 1$. For the dual structure of $\mathcal{E}_{crf}$, we construct an r-round cipher $E_r \in \mathcal{E}_{crf}$ such that $cor((\delta_0, \delta_1) \cdot x - (\delta_r, \delta_{r+1}) \cdot E_r(x)) \neq 0$. If $r = 1$, let $S_1(x) = a_1 x$ for $\delta_1^1 \neq 0$ and any linear transformation over $\mathbb{F}_p$ otherwise. Then all operations in $E_1 \in \mathcal{E}_{crf}$ are linear over $\mathbb{F}_p$, which implies that there exists an affine transformation $L_1(x) = A_1 x + B_1$, where $x \in \mathbb{F}_p^t$, A_1 is a $t \times t$ matrix over $\mathbb{F}_p$ and B_1 is a t-dimensional vector over $\mathbb{F}_p$, such that $E_1(x) = L_1 x$ and with

$$cor((\delta_1^1, \cdots, \delta_1^t) \cdot x - (\delta_2^1, \cdots, \delta_2^t) \cdot E_1(x)) = 1.$$

  Assume that we have $E_{r-1}(x) = L_{r-1} x = A_{r-1} x + B_{r-1}$ where with A_{r-1} is a $t \times t$ matrix over $\mathbb{F}_p$ and B_{r-1} is a t-dimensional vector over $\mathbb{F}_p$ such that $cor((\delta_1^1, \cdots, \delta_1^t) \cdot x - (\delta_r^1, \cdots, \delta_r^t) \cdot E_{r-1}(x)) = 1$.

  We then define $S_r(x)$ in the r-th and get $E_r(x) = L_r x = A_r x + B_r$ where with A_r is a $t \times t$ matrix over $\mathbb{F}_p$ and B_r is a t-dimensional vector over $\mathbb{F}_p$ such that $cor((\delta_1^1, \cdots, \delta_1^t) \cdot x - (\delta_{r+1}^1, \cdots, \delta_{r+1}^t) \cdot E_r(x)) = 1$. Thus, we have $cor((\delta_1^1, \cdots, \delta_1^t) \cdot x - (\delta_{r+1}^1, \cdots, \delta_{r+1}^t) \cdot E_r(x)) \neq 0$.

Considering these two parts, the proof can be finished by a similar way in Theorem 1. □

B Proof of Lemma 4

Proof. Let $C^2 = \sum_{a \in A} |cor(a^T \cdot x - b^T \cdot F(x))|^2$, then we have

$$\begin{aligned}
C^2 &= \sum_{a \in A} \frac{1}{p^t} \sum_{x \in \mathbb{F}_p^t} e^{\frac{2\pi i}{p}(a^T \cdot x - b^T \cdot F(x))} \times \frac{1}{p^t} \overline{\sum_{\lambda \in \mathbb{F}_p^t} e^{\frac{2\pi i}{p}(a^T \cdot \lambda - b^T \cdot F(\lambda))}} \\
&= \sum_{a \in A} \frac{1}{p^t} \sum_{x \in \mathbb{F}_p^t} e^{\frac{2\pi i}{p}(a^T \cdot x - b^T \cdot F(x))} \times \frac{1}{p^t} \sum_{\lambda \in \mathbb{F}_p^t} e^{\frac{-2\pi i}{p}(a^T \cdot \lambda - b^T \cdot F(\lambda))} \\
&= \frac{1}{p^{2t}} \sum_{x \in \mathbb{F}_p^t} \sum_{\lambda \in \mathbb{F}_p^t} e^{\frac{2\pi i}{p} b^T \cdot F(\lambda) - b^T \cdot F(x)} \sum_{a \in A} e^{\frac{2\pi i}{p} a^T \cdot (x - \lambda)}
\end{aligned}$$

Now, let $\theta = x - \lambda$

$$\begin{aligned} C^2 &= \frac{1}{p^{2t}} \sum_{\lambda+\theta \in \mathbb{F}_p^t} \sum_{\lambda \in \mathbb{F}_p^t} e^{\frac{2\pi i}{p} b^T \cdot F(\lambda) - b^T \cdot F(\lambda+\theta)} \sum_{a \in A} e^{\frac{2\pi i}{p} a^T \cdot \theta} \\ &= \frac{1}{p^{2t}} \sum_{\theta \in A^\perp} \sum_{\lambda \in \mathbb{F}_p^t} e^{\frac{2\pi i}{p} b^T \cdot F(\lambda) - b^T \cdot F(\lambda+\theta)} |A| \\ &= \frac{1}{p^t} \sum_{\lambda \in \mathbb{F}_p^t} e^{\frac{2\pi i}{p} (b^T \cdot F(\lambda))} \frac{1}{|A^\perp|} \sum_{\theta \in A^\perp} e^{\frac{2\pi i}{p} (-b^T \cdot F(\lambda+\theta))} \\ &= \frac{1}{p^t} \sum_{\lambda \in \mathbb{F}_p^t} e^{\frac{2\pi i}{p} (b^T \cdot F(\lambda))} cor(-b^T \cdot G_\lambda(\theta)). \end{aligned}$$

As claimed. □

C Experiments of ZC, INT of GMiMC

We now present the details of the experiments on GMiMC's ZC and INT over $\mathbb{F}_p$ as below.

For GMiMC$_{\texttt{erf}}$:

- ZC for 9-round GMiMC$_{\texttt{erf}}$-$(p = 11, t = 4)$: We check the following ZC

  $$(1, 1, 1, -2) \overset{\text{9R-ZC}}{\nrightarrow} (-4, 2, 2, 2),$$

  where its correlation over $\mathbb{F}_p$ is zero.

- ZC for 11-round GMiMC$_{\texttt{erf}}$-$(p = 11, t = 4)$: We check the following ZC

  $$(1, 1, 1, -2) \overset{\text{11R-ZC}}{\nrightarrow} (-2, 1, 1, 1),$$

  where its correlation over $\mathbb{F}_p$ is zero.

- LC for any number of rounds GMiMC$_{\texttt{erf}}$-$(p = 5, t = 6)$: We check the following LC

  $$(1, 1, 1, 1, 1, 1) \overset{\text{Any Round LC}}{\nrightarrow} (1, 1, 1, 1, 1, 1),$$

  where its linear probability is 1. This trail also leads to any number of rounds ZCs.

- INT for 9-round GMiMC$_{\texttt{erf}}$-$(p = 11, t = 4)$: We check the 9 rounds INT with input $\{(x_0, x_1, 2x_3 - x_0 - x_1, x_3) | (x_0, x_1, x_3) \in \mathbb{F}_p^3\}$, then the sum $-2y_0 + y_1 + y_2 + y_3$ of the output (y_0, y_1, y_2, y_3) is balanced.

- INT for any number of rounds GMiMC$_{\texttt{erf}}$-$(p = 5, t = 6)$: We check the any number of rounds INT with input $\{(x_0, x_1, x_2, x_3, x_4, x_0 + x_1 + x_2 + x_3 + x_4) | (x_0, x_1, x_2, x_3, x_4) \in \mathbb{F}_p^5\}$, then the sum y_0 of the output $(y_0, y_1, y_2, y_3, y_4, y_5)$ is balanced.

  For another any number of rounds INT with input $\{(0, x_1, 1, 2, 3, 4) | x_1 \in \mathbb{F}_p\}$, the sum $y_0 + y_1 + y_2 + y_3 + y_4 + y_5$ of the output $(y_0, y_1, y_2, y_3, y_4, y_5)$ is balanced.

For GMiMC$_{\texttt{crf}}$:

- ZC for 9-round GMiMC$_{\texttt{crf}}$-$(p = 11, t = 4)$: We check the following ZC

  $$(0, 0, 0, 2) \overset{\text{9R-ZC}}{\nrightarrow} (1, 0, 0, 0),$$

  where its correlation over $\mathbb{F}_p$ is zero.

- ZC for 11-round GMiMC$_{\texttt{crf}}$-$(p = 11, t = 4)$: We check the following ZC

$$(0, 0, 0, 2) \overset{\text{11R-ZC}}{\nrightarrow} (2, 0, 0, 0),$$

  where its correlation over $\mathbb{F}_p$ is zero.

- ZC for any number of rounds GMiMC$_{\texttt{crf}}$-$(p = 5, t = 6)$: We check the following ZC

$$(0, 0, 0, 0, 0, -1) \overset{\text{Any Round ZC}}{\nrightarrow} (1, 0, 0, 0, 0, 0),$$

  where its correlation over $\mathbb{F}_p$ is zero.

- INT for 9-round GMiMC$_{\texttt{crf}}$-$(p = 11, t = 4)$: We check 9 rounds INT with input $\{(x_0, x_1, x_2, 1) | (x_0, x_1, x_2) \in \mathbb{F}_p^3\}$, the sum y_0 of the output (y_0, y_1, y_2, y_3) is balanced.

For GMiMC$_{\texttt{Nyb}}$:

- ZC for 7-round GMiMC$_{\texttt{Nyb}}$-$(p = 11, t = 4)$: We check the following ZC

$$(2, 0, 0, 0) \overset{\text{7R-ZC}}{\nrightarrow} (0, 1, 0, 0),$$

  where its correlation over $\mathbb{F}_p$ is zero.

- ZC for 9-round GMiMC$_{\texttt{Nyb}}$-$(p = 11, t = 4)$: We check the following ZC

$$(0, 0, 1, 0) \overset{\text{9R-ZC}}{\nrightarrow} (0, 1, 0, 0),$$

  where its correlation over $\mathbb{F}_p$ is zero.

- INT for 7-round GMiMC$_{\texttt{Nyb}}$-$(p = 11, t = 4)$: We check 7 rounds INT with input $\{(1, x_1, x_2, x_3) | (x_1, x_2, x_3) \in \mathbb{F}_p^3\}$, the sum y_1 of the output (y_0, y_1, y_2, y_3) is balanced.

For GMiMC$_{\texttt{mrf}}$:

- ZC for 7-round GMiMC$_{\texttt{mrf}}$-$(p = 11, t = 4)$: We check the following ZC

$$(0, 2, 0, 0) \overset{\text{7R-ZC}}{\nrightarrow} (0, 0, 0, 1),$$

  where its correlation over $\mathbb{F}_p$ is zero.

- ZC for 9-round GMiMC$_{\texttt{mrf}}$-$(p = 11, t = 4)$: We check the following ZC

$$(0, 1, 0, 0) \overset{\text{9R-ZC}}{\nrightarrow} (0, 0, 0, 1),$$

  where its correlation over $\mathbb{F}_p$ is zero.

- INT for 7-round GMiMC$_{\texttt{mrf}}$-$(p = 11, t = 4)$: We check 7 rounds INT with input $\{(x_0, 1, x_2, x_3) | (x_0, x_2, x_3) \in \mathbb{F}_p^3\}$, the sum y_3 of the output (y_0, y_1, y_2, y_3) is balanced.